

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:56:55 / Search time 66 Seconds  
(without alignments)  
4152.592 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKREDFKVGNLGK.....KLQCLSLILMFSPNPNH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp19808:\*\n2: geneseqp19908:\*\n3: geneseqp20008:\*\n4: geneseqp20018:\*\n5: geneseqp20028:\*\n6: geneseqp20038:\*\n7: geneseqp20048:\*\n8: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5075	99.9	970	AAW78833	AAW78833 Human pro
2	5075	99.9	970	AAW78833	AAW78833 Human pro
3	5075	99.9	980	AAW79817	AAW79817 Human pro
4	5075	99.9	980	AAW41030	AAW41030 Human pro
5	3927.5	77.3	925	AAW92177	AAW92177 SAK-a ser
6	3927.5	77.3	925	ABBS7273	ABBS7273 Mouse lsc
7	1926	37.9	464	AAW92214	AAW92214 SAK-b ser
8	1883.5	37.1	416	AAW92176	AAW92176 SAK serin
9	1341.5	26.4	769	ABBS3067	ABBS3067 Drosophila
10	599.5	11.8	576	ABBS6250	ABBS6250 Drosophila
11	590.5	11.6	607	AAW34495	AAW34495 Human PRK
12	590.5	11.6	607	ADA50174	ADA50174 Human PRK
13	577	11.4	626	AAW35805	AAW35805 Protein f
14	560.5	11.0	685	AAW88432	AAW88432 Disease a
15	560.5	11.0	685	AAW00915	AAW00915 Human ser
16	560.5	11.0	685	ABP61474	ABP61474 Human NF-
17	560.5	11.0	753	ABP41992	ABP41992 Human ova
18	547	10.8	603	AAW79317	AAW79317 Mouse pol
19	545	10.7	603	AAW79306	AAW79306 Mouse pol
20	545	10.7	603	AAW79318	AAW79318 Mouse pol
21	540	10.6	603	AAW79312	AAW79312 Mouse pol
22	540	10.6	603	AAW79320	AAW79320 Mouse pol
23	540	10.6	603	AAW79314	AAW79314 Mouse pol
24	540	10.6	603	AAW79310	AAW79310 Mouse pol
25	539	10.6	603	AAW79309	AAW79309 Mouse pol

26	539	10.6	603	AAW79316	AAW79316 Mouse pol
27	539	10.6	603	AAW79319	AAW79319 Mouse pol
28	539	10.6	603	AAW79308	AAW79308 Mouse pol
29	538	10.6	603	AAW79311	AAW79311 Mouse pol
30	537	10.6	603	AAW79315	AAW79315 Mouse pol
31	536	10.6	603	ABR48196	ABR48196 Human bla
32	533	10.5	603	AAW79313	AAW79313 Mouse pol
33	533	10.5	603	ABR92138	ABR92138 Human cer
34	528	10.4	603	AAW74620	AAW74620 Human lsc
35	526.5	10.4	603	ABBS3219	ABBS3219 Protein s
36	524	10.3	523	AAW03425	AAW03425 Wheat put
37	524	10.3	523	ABR40719	ABR40719 Triticum
38	521.5	10.3	739	ABBS3486	ABBS3486 Drosophila
39	516.5	10.2	329	AAW56690	AAW56690 Human pro
40	516.5	10.2	832	AAW68340	AAW68340 A human d
41	516.5	10.2	832	AAW93053	AAW93053 Drosophila
42	510.5	10.1	402	ABR61579	ABR61579 Human Hba
43	510.5	10.1	402	ADD89969	ADD89969 Human can
44	510	10.0	1349	AAW70854	AAW70854 C albican
45	508.5	10.0	403	AAW18084	AAW18084 Human Aur

## ALIGNMENTS

RESULT 1	
AAW78833	
ID AAW78833	standard; protein; 970 AA.
AC AAW78833;	
DT 06-NOV-2001	(first entry)
XX	
XX	
DE Human protein SEQ ID NO 1495.	
XX	
XX	
KM Human; cytokine; cell proliferation; gene therapy;	
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KM tissue growth factor; immunomodulatory; cancer; leukaemia;	
KM nervous system disorder; arthritis; inflammation.	
XX	
OS Homo sapiens.	
XX	
PN W0200157190-A2.	
XX	
PD 09-AUG-2001.	
XX	
XX	
PF 05-FEB-2001; 2001WO-US004098.	
XX	
PR 03-FEB-2000; 2000US-00496914.	
PR 27-APR-2000; 2000US-00560875.	
PR 20-JUN-2000; 2000US-00598075.	
PR 19-JUL-2000; 2000US-00620325.	
PR 01-SEP-2000; 2000US-00654936.	
PR 15-SEP-2000; 2000US-00663561.	
PR 20-OCT-2000; 2000US-00693325.	
PR 30-NOV-2000; 2000US-00728422.	
XX	
PA (HYSE-) HYSEQ INC.	
XX	
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI Me Y, Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW;	
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX	
XX WPI, 2001-476283/51.	
DR N-PSDB; AAK51966.	
XX	
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT in diagnosis and gene therapy.	
XX	
PS Claim 20, Page 3782-3784; 6221pp; English.	
XX	
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the	
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity relating to	

CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAK60020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 CC  
 XX

Sequence 970 AA;

Query Match 99.9%; Score 5075; DB 4; Length 970;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCGEKIEPKVGNLKGSPAGVYRAESIHFGLEVALIMIDKAMYKAGVQVQNE 60
DB 1 MATCGEKIEPKVGNLKGSPAGVYRAESIHFGLEVALIMIDKAMYKAGVQVQNE 60
QY 61 VKIHCOLGPSILLEYNFEDSNVYLYLEMCHNGEMNRVYKQVKEPSEARHFMQI 120
DB 61 VKIHCOLGPSILLEYNFEDSNVYLYLEMCHNGEMNRVYKQVKEPSEARHFMQI 120
QY 121 ITGWLALHSHGILHRDITLSMLITRRNNIKIADFGLATOLKMPHEKHVTLCTPYNISP 180
DB 121 ITGWLALHSHGILHRDITLSMLITRRNNIKIADFGLATOLKMPHEKHVTLCTPYNISP 180
QY 181 EIATRSAGHGESDVWSIGCKEFTYLLIGRPEDDTVTNLTANKVLAIDYEMPSTLSIAYD 240
DB 181 EIATRSAGHGESDVWSIGCKEFTYLLIGRPEDDTVTNLTANKVLAIDYEMPSTLSIAYD 240
QY 241 LIHQLRNPADRLSLSSVLDHPPMSNNSSTKSDLGTVEDSIDSGHATITATITASSST 300
DB 241 LIHQLRNPADRLSLSSVLDHPPMSNNSSTKSDLGTVEDSIDSGHATITATITASSST 300
QY 301 SISGSLFDRKRLIIGOLFENKMTVPFNKNS-TDPSSSGGNSGYTQWGOETNSGGRV 360
DB 301 SISGSLFDRKRLIIGOLFENKMTVPFNKNS-TDPSSSGGNSGYTQWGOETNSGGRV 360
QY 361 IQDAERPHSRVYLRAYSSDRSGTNSQSAQTYMERGSAEMLSVSKSGSGGSEERY 420
DB 361 IQDAERPHSRVYLRAYSSDRSGTNSQSAQTYMERGSAEMLSVSKSGSGGSEERY 420
QY 421 SPTDNNANFNPEKETSQSGFERPDNNQALSNHLCGKTPPPADPTPQETVQOMF 480
DB 421 SPTDNNANFNPEKETSQSGFERPDNNQALSNHLCGKTPPPADPTPQETVQOMF 480
QY 481 GNLQINAHLRKTYEYDVISPNRDPQHPDLOKDTSKNMTDTPKVKNSDASNNASHVQO 540
DB 481 GNLQINAHLRKTYEYDVISPNRDPQHPDLOKDTSKNMTDTPKVKNSDASNNASHVQO 540
QY 541 NTKMKWTALHSPKPEIIQOECVFGSDPLSHOSKTRGMEPPWGYONRTLASISPLVAHRLK 600
DB 541 NTKMKWTALHSPKPEIIQOECVFGSDPLSHOSKTRGMEPPWGYONRTLASISPLVAHRLK 600
QY 601 PIROKTKKAVVSIIDSEECVELVEKVASOEYVEVLOISDGMTITITYNGRGPPLA 660
DB 601 PIROKTKKAVVSIIDSEECVELVEKVASOEYVEVLOISDGMTITITYNGRGPPLA 660
QY 661 DRPSPPTNISRYSNDNPEKWKRYQYASRFVQLVRSKSPKITVFRYAKCLIMENSPG 720
DB 661 DRPSPPTNISRYSNDNPEKWKRYQYASRFVQLVRSKSPKITVFRYAKCLIMENSPG 720
QY 721 ADPEWVFDGKVIKHTEDFIOVIEKTKSYTLKSESEVNSKKEIKMYMDHANEGHRL 780
DB 721 ADPEWVFDGKVIKHTEDFIOVIEKTKSYTLKSESEVNSKKEIKMYMDHANEGHRL 780
QY 781 ALBSIISSEERKTSAPFPPIIIGRKPGSTSSPPALSPPSVDSNVTTRDRASFNRMVH 840
DB 781 ALBSIISSEERKTSAPFPPIIIGRKPGSTSSPPALSPPSVDSNVTTRDRASFNRMVH 840

```

```

QY 841 SAASPTOAPILNPSMTNBSGLTTTASGTDISSNLSKCLPKSAQLKSVFKNVGMAT 900
DB 841 SAASPTOAPILNPSMTNBSGLTTTASGTDISSNLSKCLPKSAQLKSVFKNVGMAT 900
QY 901 QLTSGAVWQFNDGSLVVOAGVSSISYSPNGQTRRYGENEKLDPYIKOKLOCLSSILL 960
DB 901 QLTSGAVWQFNDGSLVVOAGVSSISYSPNGQTRRYGENEKLDPYIKOKLOCLSSILL 960
QY 961 MFSNPINPH 970
DB 961 MFSNPINPH 970

```

RESULT 2  
 AAM39244  
 ID AAM39244 standard; protein; 970 AA.  
 XX  
 AC AAM39244;  
 XX  
 DT 22-OCT-2001 (first entry)  
 DE Human polypeptide SEQ ID NO 2389.  
 XX  
 XX Human; nocotropic; immunosuppressant; cyostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW leukinaemia; thrombolytic; drug screening; arthritis; inflammation;  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX  
 PA (HYSB-) HYSBQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,  
 PI Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58400.  
 XX  
 XX Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 PT  
 PS Example 4; SEQ ID NO 2389; 10078pp; English.  
 XX  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAW42213) with nocotropic,  
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC

CC Activin/inhibin activity, chemocactive/chemokinetetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

Query Match	99.9%	Score 5075;	DB 4;	Length 970;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 969;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MATCIGEKIEBFKGNLNLGKGSFAGVYAAEISHTGLAEVAIKMIDKKAMYKAGMYOVONE	60	
Db	1	MATCIGEKIEDPFKGNLLGKGSFAGVYAAEISHTGLAEVAIKMIDKKAMYKAGMYOVONE	60	
QY	61	VKIHCOLKHPISLELYNYFEEDSNYYVYLEMCHGENMRYLKNNYKPPSENAHFHQI	120	
Db	61	VKIHCOLKHPISLELYNYFEEDSNYYVYLEMCHGENMRYLKNNYKPPSENAHFHQI	120	
QY	121	ITGWLHYHSHGILHRDLTLSNLLLTRNNNITKADGLATOLKMPHEKHYTLGGPNYISP	180	
Db	121	ITGWLHYHSHGILHRDLTLSNLLLTRNNNITKADGLATOLKMPHEKHYTLGGPNYISP	180	
QY	181	EIATRSAGHGESDWSLGCMEYTLILGRPEPDYDVKNKTLKVVYADVEMSPFSLIEAKD	240	
Db	181	EIATRSAGHGESDWSLGCMEYTLILGRPEPDYDVKNKTLKVVYADVEMSPFSLIEAKD	240	
QY	241	LIHOLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGAATISTAITASSST	300	
Db	241	LIHOLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGAATISTAITASSST	300	
QY	301	SISGSLPDKRLLIGOLPKNQVYFPKKKSTDPSSSGDGNSFYIOMQNOEFTSNSGRRV	360	
Db	301	SISGSLPDKRLLIGOLPKNQVYFPKKKSTDPSSSGDGNSFYIOMQNOEFTSNSGRRV	360	
QY	361	IQDAEERPHSYLRLRAYSSDRSGTNSGSAQKTYMERCHSAEMLSVSKRSGGGENERY	420	
Db	361	IQDAEERPHSYLRLRAYSSDRSGTNSGSAQKTYMERCHSAEMLSVSKRSGGGENERY	420	
QY	421	SPTNNNANIENFPEKETSSSSGSPERPNNQALSNHLCPGKTPPPPADPTPOTETVQOMF	480	
Db	421	SPTNNNANIENFPEKETSSSSGSPERPNNQALSNHLCPGKTPPPPADPTPOTETVQOMF	480	
QY	481	GNLGINHLARKTTEVDSISPMDFGCHDLCXDTSKNAATPDKVQNSDASDNASHSVQO	540	
Db	481	GNLGINHLARKTTEVDSISPMDFGCHDLCXDTSKNAATPDKVQNSDASDNASHSVQO	540	
QY	541	NTMKYMTALSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYONTLRSITPLVAHRLK	600	
Db	541	NTMKYMTALSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYONTLRSITPLVAHRLK	600	
QY	601	PIROKTKAAVSIIDSEEVCELYEVAISOEYKAEVLOISSDGNTITITYYPNGCHGFPLA	660	
Db	601	PIROKTKAAVSIIDSEEVCELYEVAISOEYKAEVLOISSDGNTITITYYPNGCHGFPLA	660	
QY	661	DRPSPPTDNISRSYSPDNLPEKYMRKYOYASRFVOLVRSKSPITITYFTFYAKCILMENSFG	720	
Db	661	DRPSPPTDNISRSYSPDNLPEKYMRKYOYASRFVOLVRSKSPITITYFTFYAKCILMENSFG	720	
QY	721	ADPEVWPYDGVKIHKTEDFIOVIEKTXGASYTLKSESEVNSLKEBIKMTMDHANEHRICL	780	
Db	721	ADPEVWPYDGVKIHKTEDFIOVIEKTXGASYTLKSESEVNSLKEBIKMTMDHANEHRICL	780	
QY	781	ALBSIIISEEERKTSAPFPPIIIGKPKSTSPKLSPPPSYDSNYPTRDRASFRRMWH	840	
Db	781	ALBSIIISEEERKTSAPFPPIIIGKPKSTSPKLSPPPSYDSNYPTRDRASFRRMWH	840	
QY	841	SAASPQOPIINPMSVNTVEGLTJTAGSTDISNSLDCLPKSQOLKSVYKVVGVNAT	900	
Db	841	SAASPQOPIINPMSVNTVEGLTJTAGSTDISNSLDCLPKSQOLKSVYKVVGVNAT	900	
QY	901	QUTSGAVVWVGQENDSQLVVAGVSSISYSPNGQTRRYGENEKLDPYIKQKLCIASSILL	960	

Db	901	Q L T S G A W Q P F N D S Q L V V Q A G V S S I S T S P N G Q T T R I G E N E K L P D I Y I K Q K Q C L S I L L	960
Ox	961	M P S N P T P N F H	970
Db	961	M P S N P T P N F H	970

Query Match	Best Local Similarity	Score 99.9%	Pred. No. 0	DB 4	Length 980
Matches 969	Conservative 1	Mismatches 0	Indels 0	Gaps 0	

```
QY 1 MATCIGEKIEDPKVGNLKGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGVQVONE 60
Db 11 MATCIGEKIEDPKVGNLKGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGVQVONE 70
QY 61 VKIHQQLKHPISILLYNYFEDSNVYVLYLEMCNGENMRYLKNNVKPSENEARHPHQI 120
Db 71 VKIHQQLKHPISILLYNYFEDSNVYVLYLEMCNGENMRYLKNNVKPSENEARHPHQI 130
QY 121 ITGMLYHSHGILHRDITLSNLLTRMNNIKIADPGIATQKMPHEGHTLCTGPNYISP 180
Db 131 ITGMLYHSHGILHRDITLSNLLTRMNNIKIADPGIATQKMPHEGHTLCTGPNYISP 190
QY 181 EIATRSAGHLESDVMSLGCMEFYTLIGRPEDDITVANKVTLADYEMPEFLSTIAXD 240
Db 191 EIATRSAGHLESDVMSLGCMEFYTLIGRPEDDITVANKVTLADYEMPEFLSTIAXD 250
QY 241 LIIHQLLRNPADRLSSVLDHPMSRNSSTKSDICTVEDSIDSGHATTSTAITSAST 300
Db 251 LIIHQLLRNPADRLSSVLDHPMSRNSSTKSDICTVEDSIDSGHATTSTAITSAST 310
QY 301 SLSGLFDKRRLLIGQPLPNKMTVFPKNSSTDPSSSGDGNSTYTOGNETNSGRGV 360
Db 311 SLSGLFDKRRLLIGQPLPNKMTVFPKNSSTDPSSSGDGNSTYTOGNETNSGRGV 370
QY 361 IQDAERPHSRYYLRAYSSDRSGTNSQSAKTYTWERCHSAEMLSVSKRSGGENEBRY 420
Db 371 IQDAERPHSRYYLRAYSSDRSGTNSQSAKTYTWERCHSAEMLSVSKRSGGENEBRY 430
QY 421 SPDDNNANINFPFKETSSSSSGSPERPDNNQALSNHLCEKTPPPADPTPOETTYQWF 480
Db 431 SPDDNNANINFPFKETSSSSSGSPERPDNNQALSNHLCEKTPPPADPTPOETTYQWF 490
QY 481 GNIQINAHLRKTEYDISISPNRDFQGHPLQKDTSKXAWTDVKKNKSDASDVAHSVQO 540
Db 491 GNIQINAHLRKTEYDISISPNRDFQGHPLQKDTSKXAWTDVKKNKSDASDVAHSVQO 550
QY 541 NTKKWTALHSKREIIQQECVRSDDLSEOSKTRGMEPPWGYQNTLRISITSPVAHRLK 600
Db 551 NTKKWTALHSKREIIQQECVRSDDLSEOSKTRGMEPPWGYQNTLRISITSPVAHRLK 610
QY 601 PIROKTKKAVSILDSSEVCEVLEKVEYAOEYKVEYLOISSDQNTITTYPPNGGGEFPA 660
Db 611 PIROKTKKAVSILDSSEVCEVLEKVEYAOEYKVEYLOISSDQNTITTYPPNGGGEFPA 670
QY 661 DRPSTPTDNISKSPNNLEPKTWKRYQVARSFVQVLRKSPKITTYFTRVAKCITLMEVSPG 720
Db 671 DRPSTPTDNISKSPNNLEPKTWKRYQVARSFVQVLRKSPKITTYFTRVAKCITLMEVSPG 730
QY 721 ADPEVWFYGVGKIKHKTEDTIOYIEKTKGKYTLKSESVNSLKEEIKMYMDANEGRICL 780
Db 731 ADPEVWFYGVGKIKHKTEDTIOYIEKTKGKYTLKSESVNSLKEEIKMYMDANEGRICL 790
QY 781 ALESIISSEERKTRSAFPFPIIIGRKPSTSSPKALSPPPVSDSNYPTRDRASFRRMWH 840
Db 791 ALESIISSEERKTRSAFPFPIIIGRKPSTSSPKALSPPPVSDSNYPTRDRASFRRMWH 850
QY 841 SAASPTQAPILNPSWMTWEGELTTTASGTDISSNSLQDCPKSAQLIKSVFVKNVWAT 900
Db 851 SAASPTQAPILNPSWMTWEGELTTTASGTDISSNSLQDCPKSAQLIKSVFVKNVWAT 910
QY 901 QLTSGAVWQFNDGSQLVQAGVSSISYTSPPNGOTTRRGNEKLPDYIKOLQCCISLTL 960
Db 911 QLTSGAVWQFNDGSQLVQAGVSSISYTSPPNGOTTRRGNEKLPDYIKOLQCCISLTL 970
QY 961 MFSNPTPNFH 970
Db 971 MFSNPTPNFH 980
```

```
RESULT 4
AAM41030
ID AAM41030 standard; protein; 980 AA.
```

```
XX AC AAM41030;
XX XX
XX 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 5961.
XX XX
XX KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX XX
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX XX
XX PF 26-DEC-2000; 2000MO-US04263.
XX XX
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX XX
XX PI Tang YT, Liu C, Aundri V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Zhou J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dirmacac RT;
XX XX
XX DR WPI; 2001-442253/47.
XX DR N-PSDB; AAI60186.
XX XX
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX PT as central nervous system injuries.
XX XX
XX PS Example 2; SEQ ID NO 5961; 10078pp; English.
XX XX
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and the
XX CC encoded polypeptides (AAM38642-AA44213) with nocotropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC C.N.S disorders. Note: The sequence data for this patent did not form
XX CC part of the printed specification
XX SQ Sequence 980 AA;
```

```
Query Match 99.9%; Score 5075; DB 4; Length 980;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 963; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MATCIGEKIEDPKVGNLKGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGVQVONE 60
Db 11 MATCIGEKIEDPKVGNLKGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGVQVONE 70
QY 61 VKIHQQLKHPISILLYNYFEDSNVYVLYLEMCNGENMRYLKNNVKPSENEARHPHQI 120
```

```

Db 71 VKHCOLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYRPFSENEARHFMHQI 130
Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHVTLCTPNYISP 180
Db 131 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHVTLCTPNYISP 190
Qy 181 EIATRSAGHLESVDWVSLGCMFYTLIGRPPFDYDVTYKNTLVADYEMPSFLSTEAD 240
Db 191 EIATRSAGHLESVDWVSLGCMFYTLIGRPPFDYDVTYKNTLVADYEMPSFLSTEAD 250
Qy 241 LTHQLRRNPADRLSLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATITSTAITSSST 300
Db 251 LTHQLRRNPADRLSLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATITSTAITSSST 310
Qy 301 STSGSLPDKRRLILGOLPKNKMTVPPKNSSTDPSSSGDGNSTYTOGNETNSGGRV 360
Db 311 STSGSLPDKRRLILGOLPKNKMTVPPKNSSTDPSSSGDGNSTYTOGNETNSGGRV 370
Qy 361 IODAEERPHSRVLRAYSSDRSGTNSQOAKTYTMRCHSAEMLSVSKRSGGGENBERY 420
Db 371 IODAEERPHSRVLRAYSSDRSGTNSQOAKTYTMRCHSAEMLSVSKRSGGGENBERY 430
Qy 421 SPTDNNANIFNFKETSSSSGSPERPDNNQALSNHLCPGKTEPPPADPTPOTETVQOMF 480
Db 431 SPTDNNANIFNFKETSSSSGSPERPDNNQALSNHLCPGKTEPPPADPTPOTETVQOMF 490
Qy 481 GNLQINAHARKTTEYDISISPNRDFQGHPDLOKDTSKNMTDTYKKNASDASNAHSVKKQ 540
Db 491 GNLQINAHARKTTEYDISISPNRDFQGHPDLOKDTSKNMTDTYKKNASDASNAHSVKKQ 550
Qy 541 NNMKWTALHSHKEIIQOECVFGSDPLSEGSKTRGMEPPGYONRTLRSTSLVLAHRLK 600
Db 551 NNMKWTALHSHKEIIQOECVFGSDPLSEGSKTRGMEPPGYONRTLRSTSLVLAHRLK 610
Qy 601 PIROKTKKAVNSILDESEVCEVLEKVEYASQEVYKEVLIQISSDNTITITYPNGRGFPPLA 660
Db 611 PIROKTKKAVNSILDESEVCEVLEKVEYASQEVYKEVLIQISSDNTITITYPNGRGFPPLA 670
Qy 661 DRPSPPTDNISRYSPDNLPKTRKQYASRFVQVLRVSKSPKITTYTRVAKCILMENSFG 720
Db 671 DRPSPPTDNISRYSPDNLPKTRKQYASRFVQVLRVSKSPKITTYTRVAKCILMENSFG 730
Qy 721 ADEPWFYGVGKTHKEDFLOVEKTKGKSYTLKSESEVNSLKEIKMYMDHANEGRICL 780
Db 731 ADEPWFYGVGKTHKEDFLOVEKTKGKSYTLKSESEVNSLKEIKMYMDHANEGRICL 790
Qy 781 ALESIISSEERKTRSAFPFIIIGRRKPGSTSPKALSPPSVDSNYPTRDRASFNRMVMH 840
Db 791 ALESIISSEERKTRSAFPFIIIGRRKPGSTSPKALSPPSVDSNYPTRDRASFNRMVMH 850
Qy 841 SAASPTOABILNFSMTNGLGLTTTASGTDISSNSLKDCLPKSAOLKSVFVKNVGMAT 900
Db 851 SAASPTOABILNFSMTNGLGLTTTASGTDISSNSLKDCLPKSAOLKSVFVKNVGMAT 910
Qy 901 QLTSGVWVOFNDGSLVQAGVSSISYSPNGQTRRYGENEGLPYIYKQKLOCLSSILL 960
Db 911 QLTSGVWVOFNDGSLVQAGVSSISYSPNGQTRRYGENEGLPYIYKQKLOCLSSILL 970
Qy 961 MFSNPTNPFH 970
Db 971 MFSNPTNPFH 980

```

RESULT 5  
 AAR92177  
 ID AAR92177 standard; protein; 925 AA.  
 AC AAR92177;  
 XX  
 DT 25-MAY-1996 (first entry)  
 XX  
 DE SAK-a serine-threonine kinase.

```

XX KW SAK-a; serine-threonine kinase; STK; agonist; antagonist;  

KW proliferative disease; cancer; tumour; antisense; transgenic animal;  

KW therapy.  

OS Mus musculus.  

XX  

FH Key Location/Qualifiers  

FT Domain 1..416  

FT /label="N-terminal domain  

FT /note="contains the kinase domain"  

FT Domain 417..925  

FT /label="C-terminal domain  

FT /note="contains 3 PEST regions"  

XX  

PN CA2150789-A.  

XX  

PD 03-DEC-1995.  

XX  

PF 01-JUN-1995; 95CA-02150789.  

XX  

PR 02-JUN-1994; 94US-00252995.  

XX  

PA (MOUN ) MOUNT SINAI HOSPITAL CORP.  

XX  

PI Dennis JM, Hefernan M, Fode C;  

XX  

DR WPI, 1996-129817/14.  

DR N-PSDB; AAT08711.  

XX  

PT Nucleic acid encoding Sak serine-threonine kinase - useful for  

PT identifying modulators potentially useful in treatment or prevention of  

PT proliferative disease.  

XX  

PS Claim 5; Page 50-55; 73pp; English.  

XX  

CC 2 Isoforms, SAK-a (AAR92177) and SAK-b (AAR92214), of a novel  

CC serine/threonine kinase are associated with mitotic and meiotic cell  

CC division and are characterized by having a kinase domain at the N-  

CC terminus and 3 PEST sequences (rich in Pro, Ser, Thr, Asp, Glu) at the C-  

CC terminus. The N-terminal regions of the 2 isoforms are identical (see  

CC AAR92176). SAK-a and SAK-b can be obtained in recombinant form by expression  

CC of encoding sequences (see AAT08711-12), and used to test for inhibitory  

CC or stimulatory effects in the treatment/diagnosis of proliferative  

CC disorders such as cancer and viral (esp. HIV) infection, or used to  

CC raise antibodies  

XX  

SQ Sequence 925 AA;  


```

Query Match 77.3%; Score 3927.5; DB 2; Length 925;  
 Best Local Similarity 78.6%; Pred. No. 1.1e-288;  
 Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;

```

Qy 1 MATCIGKEIDFVGNLAKGSFAGVYRAESIHTGLEVAIKIMDKKAMTKAGVGRVONE 60
Db 1 MAACIGERIEDFRVGNLAKGSFAGVYRAESIHTGLEVAIKIMDKKAMTKAGVGRVONE 60
Qy 61 VKHCOLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYRPFSENEARHFMHQI 120
Db 61 VKHCOLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYRPFSENEARHFMHQI 120
Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHVTLCTPNYISP 180
Db 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHVTLCTPNYISP 180
Qy 181 EIATRSAGHLESVDWVSLGCMFYTLIGRPPFDYDVTYKNTLVADYEMPSFLSTEAD 240
Db 181 EIATRSAGHLESVDWVSLGCMFYTLIGRPPFDYDVTYKNTLVADYEMPSFLSTEAD 240
Qy 241 LTHQLRRNPADRLSLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATITSTAITSSST 300
Db 241 LTHQLRRNPADRLSLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATITSTAITSSST 300

```

```

QY 301 SSGSLPDKRLLIGPLPNKMTVPFNKKSSTDFSSSGDGSFYTQWGN--QETSNSGRG 358
DB 301 SLGSLLD-RRLVGGPLPNKITVFOKNKSSDF--SSGDSNFTQWGNQEGANSRG 358
QY 359 RVIQDAERPHSRILRAAYSSDSGTSNSOGAKTYTMRCHSAEMLSVKRSGGGENEE 418
DB 359 RVIDAERPHSRILRAHSSDRASPEN-OSRAKTYVERCHSEMLSKRRS----- 410
QY 419 RYSPDNNANIFNPFKEKTSSSSGSFRRPNNOALSNHLCPGKTPPPADPTQETVQ 478
DB 411 -----LBNQHSNNHCLGKTPPPADPTQETVQ 442
QY 479 WFGNLQINAHLRKTEYDSISPNRDFQHPDLQKTSKNAWTDYKVKNSDASDNASVYK 538
DB 443 WFGNLQINAHLRKTEYDSISPNRDFQHPDLQ-DTLRNAMWTDTRASKNADTSANVAHV 501
QY 539 QQNTMKMTALHSKPELIIQOECVFGSDPLSEQSKTGMPEPMGYQNTLRSTISPLVAHR 598
DB 502 QLSAMKTYMSAHHKPEVMPQEP--GLPHSBQSKNRMESTLGYOKPTLRSTISPLIAHR 559
QY 599 LKPIROKTKKAVSILDSBEVCVELYKVASQEVYKAVLQISSDGMTITITYPNNGRGP 658
DB 560 LKPIROKTKKAVSILDSBEVCVELYKVASQEVYKAVLQISSDGMTITITYPNNGRGP 619
QY 659 LADRPSPDNTSRYSFDNLPEKYMRYOYASRFVLVRSKSPKITTYFTRAKCIIEMNS 718
DB 620 LADRPSPDNTSRYSFDNLPEKYMRYOYASRFVLVRSKSPKITTYFTRAKCIIEMNS 679
QY 719 PGADPEVWYDGYKHKTEDFLOVIEKTKSYTLKSESVNSLKEIKYMTMHANGHRI 778
DB 680 PGADPEVWYDGYKHKTEDFLOVIEKTKSYTLKSESVNSLKEIKYMTMHANGHRI 739
QY 779 CLALESIIEBERKTSAPFPFIIIGRKDGSTSPALSPPSVDNSYPTDRASFNRYV 838
DB 740 CLALESIIEBERKTSAPFPFIIIGRKDGSTSPALSPPSVDNSYPTDRASFNRYV 798
QY 839 MMSAASPQAPILNPSMTNEGLTTLTASGDISNSLKDCLPKSAQLLKSVFVNVGM 898
DB 799 VNSAASPQAPILNPSMTNEGLTTLTASGDISNSLKDCLPKSAQLLKSVFVNVGM 853
QY 899 ATQLTSGAVWVQNDGSLVQAVSSISYTSNPGQTRTYGENEQLPDYIKKLOCLSI 958
DB 854 ATQLTSGAVWVQNDGSLVQAVSSISYTSNPGQTRTYGENEQLPDYIKKLOCLSI 913
QY 959 LLMFSNPTNPF 969
DB 914 LLMFSNPTNPF 924

```

## RESULT 6

```

ID ABB57273 standard; protein; 925 AA.
AC ABB57273;
XX 07-MAR-2002 (first entry)
DE Mouse ischaemic condition related protein sequence SEQ ID NO:766.
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasoospastic ischaemia; ischaemic condition; ischaemic disease.
OS Mus musculus.
XX WO200188188-A2.
XX 22-NOV-2001.
XX 18-MAY-2001; 2001WO-JP004192.
XX 18-MAY-2000; 2000JP-00145977.
XX (UUNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

```

```

XX Ishikawa K, Arai S, Takahashi Y, Nagata T, Ishii Y;
PI WPI; 2002-03473/04.
XX N-PSDB; AB199713.
DR Examining the ischaemic condition (e.g. occlusive ischaemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX Claim 2; Page 1893-1897; 2690pp; English.
PS The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasoospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX Sequence 925 AA;
SQ
Query Match 77.3%; Score 3927.5; DB 5; Length 925;
Best Local Similarity 78.6%; Pred. No. 1.1e-288;
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;
QY 1 MATCGEKIEDPKGNLIGKSPAGVYRAESIHTGLEVALIAMDKKAYKAGVQVONE 60
DB 1 MAACIGERIEDFKGNLIGKSPAGVYRAESIHTGLEVALIAMDKKAYKAGVQVONE 60
QY VKIHQULHPSVLELYNFEDSNVYLVLEMGHNGEMRYLKNVYKPESEEAHFHQI 120
DB VKIHQULHPSVLELYNFEDSNVYLVLEMGHNGEMRYLKNVYKPESEEAHFHQI 120
QY ITGMVLYHSHGILHRDLTSLNMLLTRNNIKIADFGLATQLMKPEKHYYTLCTPNYISP 180
DB ITGMVLYHSHGILHRDLTSLNMLLTRNNIKIADFGLATQLMKPEKHYYTLCTPNYISP 180
QY 181 EIATRSAGLESQVWSLQCMFYTLIGRPPDPTTVKNTLVADYEMPSLIEAKD 240
DB 181 EIATRSAGLESQVWSLQCMFYTLIGRPPDPTTVKNTLVADYEMPSLIEAKD 240
QY 241 LIHOLLRNPPADRLSLSSVLDHPMSRNSSTKSKDLGTVEDSINDGHATTSTATTASSST 300
DB 241 LIHOLLRNPPADRLSLSSVLDHPMSRNSSTKSKDLGTVEDSINDGHATTSTATTASSST 300
QY 301 SSGSLPDKRLLIGPLPNKMTVPFNKKSSTDFSSSGDGSFYTQWGN--QETSNSGRG 358
DB 301 SLGSLLD-RRLVGGPLPNKITVFOKNKSSDF--SSGDSNFTQWGNQEGANSRG 358
QY 359 RVIQDAERPHSRILRAAYSSDSGTSNSOGAKTYTMRCHSAEMLSVKRSGGGENEE 418
DB 359 RVIDAERPHSRILRAHSSDRASPEN-OSRAKTYVERCHSEMLSKRRS----- 410
QY 419 RYSPDNNANIFNPFKEKTSSSSGSFRRPNNOALSNHLCPGKTPPPADPTQETVQ 478
DB 411 -----LBNQHSNNHCLGKTPPPADPTQETVQ 442
QY 479 WFGNLQINAHLRKTEYDSISPNRDFQHPDLQKTSKNAWTDYKVKNSDASDNASVYK 538
DB 443 WFGNLQINAHLRKTEYDSISPNRDFQHPDLQ-DTLRNAMWTDTRASKNADTSANVAHV 501
QY 539 QQNTMKMTALHSKPELIIQOECVFGSDPLSEQSKTGMPEPMGYQNTLRSTISPLVAHR 598
DB 502 QLSAMKTYMSAHHKPEVMPQEP--GLPHSBQSKNRMESTLGYOKPTLRSTISPLIAHR 559

```

```

QY 599 LKPIROKTKKAVVSIIDSEVCELVKRYASOBYKRVLOISSDGMTITTYENGGRGP 658
DB 560 LKPIROKTKKAVVSIIDSEVCELVKRYASOBYKRVLOISSDGMTITTYENGGRGP 619
QY 659 LADRPSPPTDNISRYSPFDNLPEKYMRYKQYASRFVLVSKSPKITYPTFRYAKCIMENS 718
DB 620 LADRPSPPTDNISRYSPFDNLPEKYMRYKQYASRFVLVSKSPKITYPTFRYAKCIMENS 679
QY 719 PGADPFVWYDGVKIKHTEDFIOVIEKTKGKSYTLKSESEVNSLKEIKMMDHANEGHRI 778
DB 680 PGADPFVWYDGVKIKHTEDFIOVIEKTKGKSYTLKSESEVNSLKEIKMMDHANEGHRI 739
QY 779 CTALSSIISEEEKTSAPPPFPIIRKKGSTSSPALSPPEVDNPTPRDASPNRV 838
DB 740 CTALSSIISEEEKTSAPPPFPIIRKKGSTSSPALSPPEVDNPTPRDASPNRV 798
QY 839 MHSASPTOPAPILNPMVMTNEGILTTTASGTDISSNSLKDCLPKSAQLKSVFVKNVGM 898
DB 799 VNSAAPTOSPGSPSTVTVEGIGHATATGTCVSSS-----LPKSAQLKSVFVKNVGM 853
QY 899 ATQLTSGAVVQFNDGSQLVQAGVSSISYTSFNGQTRRYGENEKLPTYIKQKLOCLSSI 958
DB 854 ATQLTSGAVVQFNDGSQLVQAGVSSISYTSFNGQTRRYGENEKLPTYIKQKLOCLSSI 913
QY 959 LMFSPNTPNF 969
DB 914 LMFSPNTPNF 924

```

## RESULT 7

AA92214 ID AA92214 standard; protein; 464 AA.

AA92214; 25-MAY-1996 (first entry)

25-MAY-1996 (first entry)

SAK-b serine-threonine kinase.

SAK-b; serine-threonine kinase; STK; agonist; antagonist; proliferative disease; cancer; tumour; antisense; transgenic animal; therapy.

Mus musculus.

Mus musculus.

Key

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Location/Qualifiers  
1. 416  
/label= "N-terminal domain"  
/note= "contains the kinase domain"  
417..464  
/label= "C-terminal domain"  
/note= "contains 3 PEST regions"

CA2150789-A.

03-DEC-1995.

01-JUN-1995; 95CA-02150789.

02-JUN-1994; 94US-00252995.

(MOUN ) MOUNT SINAI HOSPITAL CORP.

Dennis JM, Hefferman M, Fode C;

WPI; 1996-129817/14.

N-PSDB; AAT08712.

Nucleic acid encoding Sak serine-threonine kinase - useful for identifying modulators potentially useful in treatment or prevention of proliferative disease.

PS Claim 6; Page 58-61; 73pp; English.

XX 2 Isoforms, SAK-a (AA92177) and SAK-b (AA92214), of a novel  
CC serine/threonine kinase are associated with mitotic and meiotic cell  
CC division and are characterized by having a kinase domain at the N-  
CC terminus and 3 PEST sequences (rich in Pro, Ser, Thr, Asp, Glu) at the C-  
CC terminus. The N-terminal regions of the 2 isoforms are identical (see  
CC AA92176). SAK-a and SAK-b can be used in recombinant form by expression  
CC of encoding sequences (see AAT08711-12), and used to test for inhibitory  
CC or stimulatory effects, useful in the treatment, diagnosis of proliferative  
CC disorders such as cancer and viral (esp. HIV) infection, or used to  
CC raise antibodies

Sequence 464 AA;

Query Match 37.9%; Score 1926; DB 2; Length 464;  
Best Local Similarity 84.9%; Pred. No. 3.3e-137;  
Matches 372; Conservative 30; Mismatches 30; Indels 6; Gaps 5;

QY 1 MATCGEKIEDPRKGNLKGSPAGYRRESIHTGELVAKIMIDKAMTKAGVORVOR 60  
DB 1 MACIGERIEDPRKGNLKGSPAGYRRESIHTGELVAKIMIDKAMTKAGVORVOR 60  
QY 61 VKIHCOLKPSILIELYNEDESNVYLVLEMCHGEMNRYLKRRVFPSENEARHFMQI 120  
DB 61 VKIHCOLKPSILIELYNEDESNVYLVLEMCHGEMNRYLKRRVFPSENEARHFMQI 120  
QY 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQLMKPEKHYTLGTPYVISP 180  
DB 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQLMKPEKHYTLGTPYVISP 180  
QY 181 EIAATSAHGLSDVWSIGCMFYTLIGRPPTDPTKNTLNKRVADYMPSPLSIEARD 240  
DB 181 EIAATSAHGLSDVWSIGCMFYTLIGRPPTDPTKNTLNKRVADYMPSPLSIEARD 240  
QY 241 LIHQLRRNPADRLSISVLDHPFMSRNSSTKSKDGLVEDSIDSGHATITATSSST 300  
DB 241 LIHQLRRNPADRLSISVLDHPFMSRNSSTKSKDGLVEDSIDSGHATITATSSST 300  
QY 301 SIIGSLFDKRLILIGOPLPKNTVFPKKSSTDFSSGCGNSFTYQGN--QETNSGRG 358  
DB 301 SIIGSLFDKRLILIGOPLPKNTVFPKKSSTDFSSGCGNSFTYQGN--QETNSGRG 358  
QY 359 RVIQDAEERPHSRYLBRASPSRSGTNSGSAKTYTMRCHSAEMLSYKSGGGENE 418  
DB 359 RVIQDAEERPHSRYLBRASPSRSGTNSGSAKTYTMRCHSAEMLSYKSGGGENE 418  
QY 419 RYSPTDNANIFNFEK 436  
DB 417 RYSPTDNANIFNFEK 434

## RESULT 8

AA92176 ID AA92176 standard; protein; 416 AA.

AA92176; 25-MAY-1996 (first entry)

25-MAY-1996 (first entry)

Sak serine-threonine kinase N-terminus.

Sak; serine-threonine kinase; STK; agonist; antagonist; proliferative disease; cancer; tumour; antisense; transgenic animal; therapy.

Mus musculus.

Mus musculus.

CA2150789-A.

03-DEC-1995.

01-JUN-1995; 95CA-02150789.

```

XX 02-JUN-1994; 94US-00252995.
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX Dennis JW, Heffernan M, Fode C;
XX WPI; 1996-129817/14.
XX N-PSDB; AAT08710.
XX
XX Nucleic acid encoding Sak serine-threonine kinase - useful for
XX identifying modulators potentially useful in treatment or prevention of
XX proliferative disease.
XX
XX Claim 3; Page 46-48; 73pp; English.
XX
XX 2 Isoforms, sak-a and sak-b, of a novel serine/threonine kinase have an
XX identical N-terminal sequence (AA92176) that contains the kinase domain
XX and that shows significant homology to the polo subfamily. The C-terminal
XX sequences (each contg. 3 PEST regions) of the 2 isoforms differ (see
XX AA92177 and AA92214). Sak-a and Sak-b are associated with mitotic and
XX meiotic cell division, and may be involved in cell proliferation. They
XX can be obtd. in recombinant form by expression of encoding sequences (see
XX AAT08710-12) and used to test for inhibitory or stimulatory cpds. that
XX may be useful in the treatment/diagnosis of proliferative disorders, such
XX as cancer and viral (esp. HIV) infection
XX
XX Sequence 416 AA;
SQ
Query Match 37.1%; Score 1883.5; DB 2; Length 416;
Best Local Similarity 87.7%; Pred. No. 4.6e-134;
Matches 362; Conservative 27; Mismatches 19; Indels 5; Gaps 4;
QY 1 MATCIGEKIDPFKVNILKGKSPAGVYRAESIHTEGLVAIKMIDKKAMTKAGVORVONE 60
DB 1 MAACIGRIEDPFKGNILKGKSPAGVYRAESIHTEGLVAIKMIDKKAMTKAGVORVONE 60
QY 61 VKIHCOLKHPSTILELYNYPEDSNVYLYLVEMCHNGEMNRYLKNRYVPSSENAHFMQI 120
DB 61 VKIHCOLKHPSTILELYNYPEDSNVYLYLVEMCHNGEMNRYLKNRYVPSSENAHFMQI 120
QY 121 ITGMLYHSHGILHRDLTSLNLTIRNNMIKIADFGLATQLMPEHKKHYTLGCTPNYISP 180
DB 121 ITGMLYHSHGILHRDLTSLNLTIRNNMIKIADFGLATQLMPEHKKHYTLGCTPNYISP 180
QY 181 EATRSAGHLESVWSLGCMTFTLLIGRPEDTDVKNLTNRYLVADYEMPSFSLSEAKD 240
DB 181 EATRSAGHLESVDIVSLGCMSTYLLIGRPEDTDVKNLTNRYLVADYEMPAFLSREAQD 240
QY 241 LTHQLRRNPADRLSLSSVLDHPMSRNSSTSKDGLTVEDSIDGHAATISVAITASSST 300
DB 241 LTHQLRRNPADRLSLSSVLDHPMSRNSSTSKDGLTVEDSIDGHAATISVAITASSST 300
QY 301 SLSGSLFDRKRLILGQPLFNKMTVPKNSSTDFFSSGSGDGSFYQWGN--QETSNSGRG 358
DB 301 SLSGSLD--RLVLVQQLPKNKITVFOKNNSDF--SSGSGSNFCQWGNPEQENSRGRG 358
QY 359 RYIDAEERPHSRYLRAVSSDRSGTSNSQAKTYTMRCHSAEMLSVSKRS 411
DB 359 RYIDAEERPHSRYLRAVSSDRSGTSNSQAKTYTMRCHSAEMLSVSKRS 411
QY 359 RYIDAEERPHSRYLRAVSSDRSGTSNSQAKTYTMRCHSAEMLSVSKRS 411
DB 359 RYIDAEERPHSRYLRAVSSDRSGTSNSQAKTYTMRCHSAEMLSVSKRS 411
RESULT 9
ABB63067
ID ABB63067 standard; protein; 769 AA.
XX
XX ABB63067;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 15993.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;

```

```

KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; AB07170.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 15993; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutic and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL20511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
XX AB82072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 769 AA;
SQ
Query Match 26.4%; Score 1341.5; DB 4; Length 769;
Best Local Similarity 33.0%; Pred. No. 1.7e-92;
Matches 321; Conservative 156; Mismatches 234; Indels 261; Gaps 24;
QY 6 GEKIDEPKVNILKGKSPAGVYRAESIHTEGLVAIKMIDKKAMTKAGVORVONEKIKC 65
DB 8 GETIEDVEVHLLKGKGFATVYKARCLTHQDAIKMIDKLLQGTGLNRYVROEVEIHS 67
QY 66 QLKHPSTILELYNYPEDSNVYLYLVEMCHNGEMNRYLKNRYVPSSENAHFMQIITGML 125
DB 68 RLKHPSTVQLTYFQDANVYLYLVELAHNGELHRVNMHIAAPPTETELASILKQVAGIL 127
QY 126 YLHSHGILHRDLTSLNLTIRNNMIKIADFGLATQLMPEHKKHYTLGCTPNYISPRIATR 185
DB 128 YLHSHIMHRDLSLNLSSREHNVKIADFGLATQLMPEHKKHYTLGCTPNYISPRIATR 187
QY 186 SAHGLSDVWSLGCMTFTLLIGRPEDTDVKNLTNRYLVADYEMPSFSLSEAKDLHQL 245
DB 188 TSHGLPADVWSVQCMLYTLVGRPREPTDAVOSTLKNVMSYIMPAHLSYDAQLINKL 247
QY 246 LRANPADRLSLSSVLDHPMSRNSSTSKDGLTVEDSIDGHAATISVAITASSSTISGS 305
DB 248 LKQLPHERITLLEAVLCHPFLMKCS-----NGHSA----- 277
QY 306 LPDKRRLILGQPLFNKMTVPKNSSTD-----FSSGSGDGSFYQWGNQETSNGRGV 360
DB 278 -----PALNVFSQSMESGSGITTFPSSDRNSQOI---RSVNSGPOOV 320
QY 361 IODAEERPHSRYLRAVSSDRSGTSNSQAKTYTMRCHSAEMLSVSKRSGGGENBERY 420
DB 321 LPQIREEPKOVHKKLPY----- 337
QY 421 SPTDNNANIPNFFKKTSSSGSFERRPNNQALSNHLCPGKTPPPFPPTPTQOTETVOQWF 480

```

Db 338 -----BOTGLF-----GQASTGLAE-----NWP 356  
 QY 481 GNLIQNAHLKRTTEVDSISPNRDPQGHPLQKQTSKAMWTDVKKNSADSNASHSVKQ 540  
 Db 357 GAAKSSAFCEA----- 368  
 QY 541 NTMKYMTALHSPREIIQOECVGSDELSEOSKTRGMEPPWGYONRTLRSTSPVLAHRLK 600  
 Db 369 -----GNVPSNQASL-----KEDRISVPLANKRL 395  
 QY 601 PIRQTKKAVNIIIDSEECVCELVK--EYASQETVKEVTLQISDQNTITITTYNGRGCP 658  
 Db 396 PTRYTKAIVMSILNRGEVLEFLKFRPTYNEDRINDICRISDDGORIITIPDPGRGHP 455  
 QY 659 IADRPSS---PTDNISRYSEFDNLPKCYMRKYOYASHFVQLVSKSPKITYFTYACIM 715  
 Db 456 VAEQPPDLQIPSGDCV-YNYDNLPSKMKKTYIGARFVGLVSKTYKTYFTSLGKQIM 514  
 QY 716 ENSPGADPEVWFYDVGVKIKHT-EDPIQVLEKTKGKSYTLKSESEVNSLKEEIKMYMDHANE 774  
 Db 515 EFM--TDFEIRFYSGAKLKTPEBGLKYVDKNGMLSDVSCSESRL-----IEHGR 565  
 QY 775 GHRICLATESIIEEERKTRSAFPPTIIGRKPGSTSPKALSPPSVDSNTPTRDRAE 834  
 Db 566 CFTHCVNISNALEVAQTKNSC--FPVTIGRPRITD-----VQPARLDG--LRDTTN- 614  
 QY 835 NNMWHSASPTQAPILNBSMTNBEGLGTTTASGTDISSNLKDLPLPSAQ--LKSVF 892  
 Db 615 ---IAFSTPKSNQSI-NFSLST-----ISSTNTSDPGTNGSRSMLAHONIPKRI 665  
 QY 893 VKNVGAATQLTSGAVWQFNDGSQLV---QAGVSIATYSPNGQTRTYGENEKLPDYI 948  
 Db 666 VEIGATLTLSHGV 723  
 QY 949 KOKLOCLSSILL 960  
 Db 724 RDVGCIPMQL 735

## RESULT 10

ABB62520

ID ABB62520 standard; protein; 576 AA.

XX AC ABB62520;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polyptide seq ID NO 14352.

XX KM Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PMD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL06623.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX PS Disclosure; SEQ ID NO 14352; 21np + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB16177), expressed DNA  
 CC sequences (AB16176-AB16177) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 576 AA;

QY Query Match 11.8%; Score 599.5; DB 4; Length 576;  
 Db Best Local Similarity 34.1%; Pred. No. 2.3e-36;  
 Matches 136; Conservative 71; Mismatches 121; Indels 71; Gaps 9;

QY 2 ATCGEKIED-----FKVGNLLGKGSFAGVYRAESIHGLEYAIXMDKAMYKAGVQ 55  
 Db 9 STDIIPRLVDINOKTYKMRFFGKGPACKEIIDVEITDVPAGKIVSKLMIKINQKE 68  
 QY 56 RVQNEVKIKCQKHSISELYNYFEDSNVYVLEWCHNGENMRYLKNEVKPSENEARH 115  
 Db 69 KTAQETIHRSLNHPVIVKFNHYFEDSONIYVLELCKRSMWELHGR-KSITEPCRY 127  
 QY 116 FMHQITGMLYHSHGILHRDITLSMLLTRMMNIKIADFGIATOLKMHKHYTICGTP 175  
 Db 128 YVYQIIGVAYLHNRKIIHRDILKNTLFNDLHVYIGDFGATRIEYGERKTCIGTP 187  
 QY 176 NYISPEIATRSAGLESVDLSGMEFYTLIGRPEDTQVKNVTLNKVLADEMPSTLS 235  
 Db 188 NYIAPETLTKGHSFEVDWISGCVWYTLVGQPPETKTLKOTYSIKCKEYRVSYR 247  
 QY 236 IEAKDLIHQLRRNPADRLSLSSVLDHPMS-----RNSSTKSLDG--TYEDSI 283  
 Db 248 KPAAADVIMLQPNPESRALIQLNLFELKSKVPMFLPSCLTVAAPIGSNDITEDSM 307  
 QY 284 D-----SG-----HATISTATNASSSTISIGSLFDRKLLIGOPLENKMT 323  
 Db 308 HRRKPLMKNNGIRPDDTRLESTFLKANLHDAITASAQVC----- 345  
 QY 324 VEPKNSSTDPSSGDSNFFYQMGNGQERTSNGRGVQ 362  
 Db 346 -----RHSEDRYS--DIESLYOQLTN--LINGKPRITQ 374

## RESULT 11

ID AAE34495 standard; protein; 607 AA.

XX AC AAE34495;

XX DT 14-MAY-2003 (first entry)

XX DE Human PRK protein.

XX KM Human; pyridylpyrimidine derivative; cellular protein kinase; Scrapie;  
 XX cellular protein phosphatase; cellular signal transduction; propylaxis;  
 XX prion infection; chronic wasting disease; CWD; Creutzfeldt-Jacob disease;  
 XX CJD; transmissible mink encephalopathy; bovine spongiform encephalopathy;  
 XX TSE; BSE; Gerstmann-Strausler-Scheiner syndrome; GSS; Alpers syndrome;  
 XX fatal familial insomnia; FFI; kuru; neurodegenerative disease; nootropic;  
 XX Alzheimer's disease; PRK.

XX OS Homo sapiens.

XX PN WO200293164-A2.

XX PD 21-NOV-2002.





KW Grave's disease; lupus erythematosus; multiple sclerosis;  
 KW myasthenia gravis; osteoarthritis; osteoporosis; pancreatitis;  
 KW polycystic kidney disease; polymyositis; scleroderma;  
 KW Sjogren's syndrome; autoimmune thyroiditis; cancer; infection; trauma;  
 KW cell proliferation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Binding-site 89..96 /note="potential ATP binding site"  
 FT Peptide 506..511 /note="presumed regulatory sequence common to polo  
 FT family protein kinases"  
 FT  
 XX WO9858052-A2.  
 XX  
 XX 23-DEC-1998.  
 XX  
 XX 19-JUN-1998; 98WO-US012813.  
 XX  
 XX 19-JUN-1997; 97US-00878989.  
 XX  
 XX (INCYTE) INCYTE PHARM INC.  
 XX  
 XX Bandman O, Hillman JL, Corley NC, Guegler KJ, Lal P, Goli SK,  
 XX Shah P,  
 XX MPI; 1999-080952/07.  
 XX N-PSDB; AAX06831.  
 XX  
 XX New disease associated protein kinases - used to stimulate cell  
 XX proliferation and to treat the immune response and cancer.  
 XX  
 XX Claim 1; Page 54-56; 93pp; English.  
 XX  
 CC This is the amino acid sequence of human disease associated protein  
 CC kinase DAPK-1, as deduced from a consensus sequence (see AAX06831) of  
 CC overlapping cDNA clones from libraries which are immortalised or  
 CC cancerous and show inflammatory or immune responses. DAPK-1 shows 53%  
 CC homology to human proliferation-related protein kinase PRK (GI 1488263).  
 CC The invention provides DAPK-1 to DAPK-7 polypeptides (see AAW8432-38)  
 CC and cDNA clones encoding them (see AAX06831-36 and AAX06882), as well as  
 CC expression vectors, host cells, agonists, antagonists and antibodies. The  
 CC invention further provides uses of such products in the diagnosis,  
 CC prevention and treatment of diseases associated with cell proliferation,  
 CC especially cancer or an immune response (claimed). Conditions that may be  
 CC treated include adult respiratory distress syndrome, allergies, asthma,  
 CC arteriosclerosis, bronchitis, emphysema, hyper eosinophilia, myocardial or  
 CC pericardial inflammation, rheumatoid arthritis, Addison's disease, AIDS,  
 CC haemophilia, atherosclerosis, various diseases of the digestive system,  
 CC atopic dermatitis, dermatomyositis, diabetes mellitus,  
 CC glomerulonephritis, gout, Grave's disease, lupus erythematosus, multiple  
 CC sclerosis, myasthenia gravis, osteoarthritis, osteoporosis, pancreatitis,  
 CC polycystic kidney disease, polymyositis, scleroderma, Sjogren's  
 CC syndrome, autoimmune thyroiditis, complications of cancer, extracorporeal  
 CC circulation, viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections, and trauma (disclosed)  
 CC  
 XX Sequence 685 AA;  
 XX  
 XX Query Match 11.0%; Score 560.5; DB 2; Length 685;  
 XX Best Local Similarity 41.6%; Pred. No. 2.8e-33;  
 XX Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;  
 QY 15 GNLGKGSFAGVYRAESIHITGLVAIKMDKKAMKAMQVONVEYKIHQLKHPSTLE 74  
 DB 85 GKVVGKGGFACQCYMTDLTNKKVYAAKIIPIHSRVAKQPHQREKIDKEITLRIIHHKHVVQ 144  
 QY 75 LYNTFPEBSNYVYVLEWCHNENRKYLNKRYKPSSENAHFHMHQITGMUYLHSHGILH 134  
 DB 145 FYHYFEDENKITYILLEYCSRSRSMHILTKAR-KVLTPEVRVRYTLNQLVSGKTYLHEQELIH 203

QY 135 RDLTISNLLITRRNNIKIADFGLATQKMPHEKHYYTLGTPNTYISPEIATRSAGHLESVD 194  
 DB 204 RDLTGKGFPIENAMELKVGFGAARLEPLEHRRRTICGTPNYLSPEVLNKGHCESDI 263  
 QY 195 WSLGCMPTTLLIGRPPTDVTIVKTKLVKAVLVADYEMSPSLSEAKDILHOLLRRNPADRL 254  
 DB 264 WALGCVMTYMLGRPPETTNLAKETRYCIRREARYTSPSLAPAKHLIASNLSKNPDRP 323  
 QY 255 SLSSVLDPHF 264  
 DB 324 SLDDIIRHDF 333  
 RESULT 15  
 AAY00915  
 ID AAY00915 standard; protein; 685 AA.  
 XX  
 XX AAY00915;  
 XX  
 XX 28-MAY-1999 (first entry)  
 XX  
 XX Human serum inducible kinase.  
 XX  
 KW Serum inducible kinase; SNK protein; human; proliferative disease;  
 KW leukaemia; solid tumour cancer; metastasis; chronic inflammatory;  
 KW psoriasis; rheumatoid arthritis; proliferative cardiovascular disease;  
 KW restenosis; ocular disorder; diabetic retinopathy; haemangioma;  
 KW benign hyperproliferative disease; diagnosis.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO9909146-A1.  
 XX  
 XX 25-FEB-1999.  
 XX  
 XX 20-AUG-1998; 98WO-US017248.  
 XX  
 XX 20-AUG-1997; 97US-0056112P.  
 XX  
 XX (SMTK) SMITHKLINE BEECHAM CORP.  
 XX  
 XX Anderson KM, Jackson JR, Hanbury MJ, Nerurkar SS, Roshak AK,  
 XX Bouzyk M,  
 XX MPI; 1999-181027/15.  
 XX N-PSDB; AAX27227.  
 XX  
 XX New serum inducible kinase (SNK) polypeptides and polynucleotides -  
 XX useful for treating proliferative diseases.  
 XX  
 XX Claim 1; Page 39-40; 41pp; English.  
 XX  
 CC This sequence is a human serum inducible kinase (SNK) of the invention.  
 CC The invention relates to diagnostic assays or kits for detecting diseases  
 CC associated with inappropriate SNK activity or levels. Disease states that  
 CC can be diagnosed include proliferative diseases such as leukemia, solid  
 CC tumour cancers and metastases, chronic inflammatory proliferative  
 CC diseases such as psoriasis and rheumatoid arthritis, proliferative  
 CC cardiovascular diseases such as restenosis, proliferative ocular  
 CC disorders such as diabetic retinopathy and benign hyperproliferative  
 CC diseases such as haemangiomas. The polynucleotides can be used as  
 CC hybridisation probes for cDNA and genomic DNA or as primers for a nucleic  
 CC acid amplification (PCR) reaction, to isolate full-length cDNAs and  
 CC genomic clones encoding polypeptides of this invention and to isolate  
 CC cDNA and genomic clones of other genes which have a high sequence  
 CC similarity to the SNK coding sequence. The differences between cDNA and  
 CC genomic sequences can be observed and therefore mutations detected. Any  
 CC mutations may then be attributed to likely causative agents of disease.  
 CC The nucleotide sequences are also useful for chromosome identification  
 CC  
 XX Sequence 685 AA;  
 XX  
 XX Query Match 11.0%; Score 560.5; DB 2; Length 685;



**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 13:58:47 ; Search time 1133 Seconds  
(without alignments)  
11340.902 Million cell updates/sec

Title: US-10-026-021-1  
Perfect score: 2913  
Sequence: 1 atgsgactcgcgcggga.....cgacctcaatttcattga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hhg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sce:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_pal:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sce:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hhg\_hum:\*  
31: em\_hhg\_in:\*  
32: em\_hhg\_other:\*  
33: em\_hhg\_mus:\*  
34: em\_hhg\_pin:\*  
35: em\_hhg\_rod:\*  
36: em\_hhg\_mam:\*  
37: em\_hhg\_vrt:\*  
38: em\_sv:\*  
39: em\_hhgo\_hum:\*  
40: em\_hhgo\_mus:\*  
41: em\_hhgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3225	9 BC036023	BC036023 Homo sapi
2	2911.4	99.9	3937	9 AR338789	AR338789 Sequence
3	2906.6	99.8	3092	9 HSSTPKSAK	Y13115 Homo sapien
4	2905	99.7	3331	9 AB006972	AB006972 Homo sapi
5	1879	64.5	3447	6 AR083569	AR083569 Sequence
6	1879	64.5	3447	6 I56871	I56871 Sequence 3
7	1879	64.5	3447	6 AX306014	AX306014 Sequence
8	1879	64.5	3447	10 MMSAKA	L29479 Mus muscu
9	1874.2	64.3	3473	10 BC026785	BC026785 Mus muscu
10	1724.6	59.2	3177	10 BC057940	BC057940 Mus muscu
11	1064.2	36.5	3554	5 BC060363	BC060363 Xenopus 1
12	1025	35.2	57057	9 AC107053	AC107053 Homo sapi
13	1025	35.2	15342	2 AL591646	AL591646 Homo sapi
14	970.4	33.3	1600	6 AR083570	AR083570 Sequence
15	970.4	33.3	1600	6 I56872	I56872 Sequence 5
16	970.4	33.3	1600	10 MMSAKB	L29480 Mus muscu
17	924.2	31.7	1453	6 AR083568	AR083568 Sequence
18	924.2	31.7	1453	6 I56870	I56870 Sequence 1
19	731.8	25.1	233181	2 AC111558	AC111558 Rattus no
20	683.4	23.5	158533	2 AC146980	AC146980 Mus muscu
21	619.6	21.3	201286	2 AC102228	AC102228 Mus muscu
22	619.6	21.3	232323	2 AC127590	AC127590 Mus muscu
23	373	12.8	2453	5 BC045434	BC045434 Danio rer
24	329.2	11.3	331	6 AX339849	AX339849 Sequence
25	313.8	10.8	972	10 BC051483	BC051483 Mus muscu
26	282.2	9.7	2574	3 AP106952	AP106952 Drosophi
27	280.6	9.6	3384	3 BT004846	BT004846 Drosophi
28	268.2	9.2	149008	9 AC007991	AC007991 Homo sapi
29	239	8.2	115884	2 AC017581	AC017581 Drosophi
30	239	8.2	168058	3 AC010049	AC010049 Drosophi
31	239	8.2	174920	3 AC010701	AC010701 Drosophi
32	239	8.2	314906	3 AE003594	AE003594 Drosophi
33	186.4	6.4	66081	2 AC083920	AC083920 Homo sapi
34	160.8	5.5	1573	3 AK113631	AK113631 Ciona int
35	154.2	5.3	381	6 AR270086	AR270086 Sequence
36	146.8	5.0	3923	9 AK128691	AK128691 Homo sapi
37	137.8	4.7	2662	5 BC042254	BC042254 Xenopus 1
38	137.8	4.7	3589	5 AP357840	AP357840 Xenopus 1
39	136	4.7	2180	9 HS085755	HS085755 Human betu
40	136	4.7	2375	9 AK098163	AK098163 Homo sapi
41	135.6	4.7	2972	9 AF059617	AF059617 Homo sapi
42	135	4.6	2381	5 XL058205	XL058205 Xenopus lae
43	135	4.6	2772	6 AX305275	AX305275 Sequence
44	135	4.6	2772	10 MMSAK	M96163 Mus muscu
45	135	4.6	2804	10 BC034513	BC034513 Mus muscu

## ALIGNMENTS

RESULT 1  
BC036023 3225 bp mRNA PRI 12-NOV-2003  
LOCUS Homo sapiens serine/threonine kinase 18, mRNA (CDNA clone MGC:33045  
DEFINITION IMAGE:5273226), complete cds.  
ACCESSION BC036023  
VERSION BC036023.1 GI:23243308  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 3225)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klauser,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLR
JOURNAL
REMARK
COMMENT
Altschul S.F., Zeeberg B., Bucosek K.H., Scheefter C.F., Bhac N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Steeple T.E., Brownstein M.J., Uediri T.B., Toshitsuki S., Carinini P., Prange C., Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J., Moyleman K.J., Malek J.A., Gunarene P.H., Richard S., Wexler K.C., Hale S., Garcia A.M., Gay L.J., Hulky S.W., Villalón D.K., Mozny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywnietki M.I., Skalska U., Smalun D.E., Scherach A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3225)
Strausberg R. Direct Submission Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk Email: gcapbs-remail.nih.gov
Tissue Procurement: Miklos Palcovics, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshitsuki and Piero Carninci (RIKEN) cDNA library Arrayed by: The I.M.A.G.E Consortium (ILNI) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.snhgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LINL ac: http://image.llnl.gov Series: IRAX Plate: 48 Row e Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21361432.
FEATURES
SOURCE
1..3225
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33045 IMAGE:5273226"
/tissue_type="Testis"
/clone_id="NIH_MGC_97"
/lab_host="DHIOB"
/note="Vector: pBluescript"
1..3225
/gene="STX18"
/note="synonym: SAK"
/db_xref="Locustid:10733"
/db_xref="MIM:605031"
243..3155
/codon_start=1
/product="serine/threonine kinase 18"
/protein_id="AAH56023.1"
/db_xref="GI:23243309"
/db_xref="Locustid:10733"
/translation="MATCIQEKIDEPYGNLLKGSPAGVRYRAISHTGCEVAIKMDKMYRAGWQRVDNVEKIHCQLKHSEIELNYPEBSRVYYIYLKCAICHGEARNTLKMKPSSENRRAPFMOIIITGMNLTLSHGLLHDLDLTNSNLLTRNNMTKLADGLATWLTKPKHKATHVLGTPTYSPELAIFRSXLGSDPVMSLGCMFYILLIGRPDDIVTGVN

		Query Match	99.9%; Score 2911.4; DB 9; Length 3225;	
		Best Local S. similarity 100.0%; Pred. No. 0;		
		Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	ATGGGACCTGTCATCGGGGAGAAATGAGAGATTTTAAAGTGGAAATCTGCTGTGTA	60	
DB	243	ATGGGACCTGTCATCGGGGAGAAATGAGAGATTTTAAAGTGGAAATCTGCTGTGTA	302	
QY	61	GGATCAATTTGCTGCTGTCTACAGAGCTGAGTCCATTCACATGGTTGGAAAGTTGCATC	120	
DB	303	GGATCAATTTGCTGCTGTCTACAGAGCTGAGTCCATTCACATGGTTGGAAAGTTGCATC	362	
QY	121	AAATGATATGATAGAAAGGCATGTATCAAAGCAGAAATGATACAGAGATGCCAAATAG	180	
DB	363	AAATGATATGATAGAAAGGCATGTATCAAAGCAGAAATGATACAGAGATGCCAAATAG	422	
QY	181	GTGAAATATGATGCCAATTTGAAACATCCTCTATCTTTGAGCTTTATATCTATTTTGA	240	
DB	423	GTGAAATATGATGCCAATTTGAAACATCCTCTATCTTTGAGCTTTATATCTATTTTGA	482	
QY	241	GATGCAATATGTGTATCTGTGTATTAAGAAATGTGCCATATGAGAAATGAAACAGTAT	300	
DB	483	GATGCAATATGTGTATCTGTGTATTAAGAAATGTGCCATATGAGAAATGAAACAGTAT	542	
QY	301	CTMAAGATGAGTGAAGAACCTTCTCAGAAATGAGAGCTGACATTCATGACACAGATC	360	
DB	543	CTMAAGATGAGTGAAGAACCTTCTCAGAAATGAGAGCTGACATTCATGACACAGATC	602	
QY	361	ATCACAGGAGTGTGTATCTTCATCTTCATGTATATCTACACCGGAGCTCACAATTTCT	420	
DB	603	ATCACAGGAGTGTGTATCTTCATCTTCATGTATATCTACACCGGAGCTCACAATTTCT	662	
QY	421	AACCTCTATGACTGTATATGAACATCAAAATGTGATTTTGGGTGGTCACTCA	480	
DB	663	AACCTCTATGACTGTATATGAACATCAAAATGTGATTTTGGGTGGTCACTCA	722	
QY	481	CTGAAATGCAATGCAAAAGCACTATATCTTATGTGAACTCTTATCTATCACTTCA	540	
DB	723	CTGAAATGCAATGCAAAAGCACTATATCTTATGTGAACTCTTATCTATCACTTCA	782	
QY	541	GAAATGCCACTGGAAGTGACATGGCTTTGAATCTGATGTTTGGTCCCTGGGCTGTATG	600	
DB	783	GAAATGCCACTGGAAGTGACATGGCTTTGAATCTGATGTTTGGTCCCTGGGCTGTATG	842	
QY	601	TTTTATACATTTACTTTATCGGAGAGACACCTTCGACATGACATGCAAGAACATTA	660	
DB	843	TTTTATACATTTACTTTATCGGAGAGACACCTTCGACATGACATGCAAGAACATTA	902	
QY	661	AATTAAGTATGTATGACAGATTAATGAATGCCATCTTTTGTGCATATAGGCAAGAC	720	

Db	903	AAATAAGTAAGTATGGCAGATTATGAAATGCAATCTTTTGTCAATAGAGGCCAAGAC	962
Oy	721	CTTATTCACCAAGTTACTTCGTAGAAATCAGACAGATCGTTTAAGCTGTCTTCAGTATTG	780
Db	963	CTTATTCACCAAGTTACTTCGTAGAAATCAGACAGATCGTTTAAGCTGTCTTCAGTATTG	1022
Oy	781	GACCATCTTTTATGTCCGAAATTCCTCAACAAAAGTAAAGATTNAGAACGTGTGAA	840
Db	1023	GACCATCTTTTATGTCCGAAATTCCTCAACAAAAGTAAAGATTNAGAACGTGTGAA	1082
Oy	841	GACCTCAATTGATAGTGGGCATGCCCAATTCATCTGCAATTAACGCTTCCTCCAGTACC	900
Db	1083	GACCTCAATTGATAGTGGGCATGCCCAATTCATCTGCAATTAACGCTTCCTCCAGTACC	1142
Oy	901	AGTATTAAGTGTAGTTTATTTGACAAAAGAGAATTGATTTGTGCAGCCACTCCCAAT	960
Db	1143	AGTATTAAGTGTAGTTTATTTGACAAAAGAGAATTGATTTGTGCAGCCACTCCCAAT	1202
Oy	961	AAAAATGACTATTTCCAAAGATAAAGTTCAACTGATTTTCTTCTTCAGAGATGGA	1020
Db	1203	AAAAATGACTATTTCCAAAGATAAAGTTCAACTGATTTTCTTCTTCAGAGATGGA	1262
Oy	1021	AACAGTTTATTAATCTCAAGTGGGGAAATCAAGAAACAGTAAATGTGGAAAGGGGAAGTA	1080
Db	1263	AACAGTTTATTAATCTCAAGTGGGGAAATCAAGAAACAGTAAATGTGGAAAGGGGAAGTA	1322
Oy	1081	ATTCAAGATGAGAAGAAAGGCCACACTCTGATACCTTCGTAGAGCTTATTCCTCTGAT	1140
Db	1323	ATTCAAGATGAGAAGAAAGGCCACACTCTGATACCTTCGTAGAGCTTATTCCTCTGAT	1382
Oy	1141	AGATCTGGCACTTCTAATATGTCAGTCTCAGCAAAACATATACATGGAACGATGTTCAC	1200
Db	1383	AGATCTGGCACTTCTAATATGTCAGTCTCAGCAAAACATATACATGGAACGATGTTCAC	1442
Oy	1201	TCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGGTGAATGAAGAGAGTAC	1260
Db	1443	TCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGGTGAATGAAGAGAGTAC	1502
Oy	1261	TCACCCACAGACAAACAATGCCAACATTTTAACTTTTAAAGAAAAGACATCCAGTAGT	1320
Db	1503	TCACCCACAGACAAACAATGCCAACATTTTAACTTTTAAAGAAAAGACATCCAGTAGT	1562
Oy	1321	TCTGGATCTTTTGAAGAAGACTGATTAACAATCAAGACACTCCCAATCACTTTGTCCAGAA	1380
Db	1563	TCTGGATCTTTTGAAGAAGACTGATTAACAATCAAGACACTCCCAATCACTTTGTCCAGAA	1622
Oy	1381	AAAACTCTTTTCCATTTTGACAGCCCGACACTCAGACTGAACCCGTACCAACATGTGTTT	1440
Db	1623	AAAACTCTTTTCCATTTTGACAGCCCGACACTCAGACTGAACCCGTACCAACATGTGTTT	1682
Oy	1441	GGGAATCTGCAATTAATATGCTCAATTAAGAAAACCTAGAAATATGACAGATACGCCA	1500
Db	1683	GGGAATCTGCAATTAATATGCTCAATTAAGAAAACCTAGAAATATGACAGATACGCCA	1742
Oy	1501	AACGGGCACTTCAGGGGCATTCAGATTGACAGAGAGACATCAAAAAATGCTCGTGACT	1560
Db	1743	AACGGGCACTTCAGGGGCATTCAGATTGACAGAGAGACATCAAAAAATGCTCGTGACT	1802
Oy	1561	GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCAATTCGTAAACAGCAA	1620
Db	1803	GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCAATTCGTAAACAGCAA	1862
Oy	1621	AATACCAATGAATATATATGACATGCACTTCAACAGTAAACCTGAGATATTCACAAGAAATGT	1680
Db	1863	AATACCAATGAATATATATGACATGCACTTCAACAGTAAACCTGAGATATTCACAAGAAATGT	1922
Oy	1681	GTTTTTGGCTCAGATCTCTTTCTGAAACAGAGCAAGACTAGGGGATATGAGCCACATGG	1740
Db	1923	GTTTTTGGCTCAGATCTCTTTCTGAAACAGAGCAAGACTAGGGGATATGAGCCACATGG	1982
Oy	1741	GGTATATCAGAAATCGTACATTAAGAGCATTAATCTCCGTTGGTTGCTCAAGGGTAAAA	1800

Db	1983	GGTTATCGAAATCGTACATTTAAGAAAGCATTAACATCTCCGTTGGTTGCTCAAGGTTAA	20142
Qy	1801	CCAAATCAGACAGAAAAACCAAAAGGCTGTGGTAGACATCTTGATTCAGAGAGGTGT	18666
Db	2043	CCAATCAGACAGAAAAACCAAAAGGCTGTGGTAGACATCTTGATTCAGAGAGGTGT	21020
Qy	1861	GTGAGCTTGTAAAGAGATATGCATCTTCAGAAATATGTGAAGAAAGTTCTCAATATCT	19220
Db	2103	GTGAGCTTGTAAAGAGATATGCATCTTCAGAAATATGTGAAGAAAGTTCTCAATATCT	21622
Qy	1921	AGTATGGAATATACGATCACTATTATTAATCAATGGTGGTAGAGGTTTTCTCTTGT	19808
Db	2163	AGTATGGAATATACGATCACTATTATTAATCAATGGTGGTAGAGGTTTTCTCTTGT	22222
Qy	1981	GATAGACCACTCTCACTACTGACAAACATCAGTAGTACAGCTTTGACAAATTTACAGAA	20404
Db	2223	GATAGACCACTCTCACTACTGACAAACATCAGTAGTACAGCTTTGACAAATTTACAGAA	22882
Qy	2041	AAATATCTGGCCGAATATATCAATATGCTTCCAGGTTTGTACAGCTTGTAAATCT	21000
Db	2283	AAATATCTGGCCGAATATATCAATATGCTTCCAGGTTTGTACAGCTTGTAAATCT	23422
Qy	2101	CCCAAAATCATCTTATTTTAAAGAAATATGCTTAAATGCAATTTGATGAGAGAAATCTCTG	21606
Db	2343	CCCAAAATCATCTTATTTTAAAGAAATATGCTTAAATGCAATTTGATGAGAGAAATCTCTG	24002
Qy	2161	GCTGATTTTGAAGTTTGTTTGTATGATGGGTAAATAACAAACAGAAATTTCAAT	22220
Db	2403	GCTGATTTTGAAGTTTGTTTGTATGATGGGTAAATAACAAACAGAAATTTCAAT	24622
Qy	2221	CAGGTGATTTGAAAGACAGGGAGTCTTACCTTTAAAGTGAATGAAATTAATATGC	22888
Db	2463	CAGGTGATTTGAAAGACAGGGAGTCTTACCTTTAAAGTGAATGAAATTAATATGC	25222
Qy	2281	TTGAAAGAGGATTAATAATGTAATGGAACATGCTAATGAGGTCATCGATTTGTTTA	23404
Db	2523	TTGAAAGAGGATTAATAATGTAATGGAACATGCTAATGAGGTCATCGATTTGTTTA	25822
Qy	2341	GCACTGGAATCCATTAATTTTCAGAAAGAGAAAGAAATCAGAGAGTGTCCCTTTTCCCA	24000
Db	2583	GCACTGGAATCCATTAATTTTCAGAAAGAGAAAGAAATCAGAGAGTGTCCCTTTTCCCA	26422
Qy	2401	ATTAATCATAGAGAGAAAACCTGTAGTAATAGTTCACTTAAGGCTTATCACTCTCTCT	24606
Db	2643	ATTAATCATAGAGAGAAAACCTGTAGTAATAGTTCACTTAAGGCTTATCACTCTCTCT	27022
Qy	2461	TTGTGTGATTCAAATATACCCCAACAGAGATAGACATCTTTCAACAGAAATGTCATGAT	25220
Db	2703	TTGTGTGATTCAAATATACCCCAACAGAGAGAGACATCTTTCAACAGAAATGTCATGAT	27622
Qy	2521	AGTGTGTCTTCTCCAAACAGAGACCAATCTCTTAATCCCTCTATGTGTTACAAATGAAGA	25808
Db	2763	AGTGTGTCTTCTCCAAACAGAGACCAATCTCTTAATCCCTCTATGTGTTACAAATGAAGA	28222
Qy	2581	CTTGTCTTCAACCTACAGCTTCTGAAACAGACAATCTTCTTAATAGCTTAAGAAATGT	26404
Db	2823	CTTGTCTTCAACCTACAGCTTCTGAAACAGACAATCTTCTTAATAGCTTAAGAAATGT	28822
Qy	2641	CTTCTTAATCAGACCACTTTTGAAATCTGTTTTGTGAAAAATGTTGTGTGGGCTACA	27000
Db	2883	CTTCTTAATCAGACCACTTTTGAAATCTGTTTTGTGAAAAATGTTGTGTGGGCTACA	29422
Qy	2701	CAGTTAATCTAGTGAGCTGTGTGGGTTCACTTAATGATGAGGTCACAGTTGGTTGTGAG	27606
Db	2943	CAGTTAATCTAGTGAGCTGTGTGGGTTCACTTAATGATGAGGTCACAGTTGGTTGTGAG	30022
Qy	2761	GCAGAGAGTGTCTTCTATCATGTTATACCTCAACAAATGCTCAACAACTAGGTATGAGAA	28220
Db	3003	GCAGAGAGTGTCTTCTATCATGTTATACCTCAACAAATGCTCAACAACTAGGTATGAGAA	30622
Qy	2821	AATGAAAAATTAACGAGACTACATCAACAGAAATTAACAGTGTCTGTCTTCAATCTCTTTG	28880
Db	3063	AATGAAAAATTAACGAGACTACATCAACAGAAATTAACAGTGTGTGTCTTCAATCTCTTTG	31222

QY 2881 ATGTTTCTAATCCGACTCTTAATTTCATTGA 2913  
DB 3123 ATGTTTCTAATCCGACTCTTAATTTCATTGA 3155

RESULT 2  
AR338789 3937 bp DNA linear PAT 17-AUG-2003  
LOCUS AR338789 Sequence 280 from patent US 6569662.  
DEFINITION AR338789  
ACCESSION AR338789.1 GI:33725646  
VERSION AR338789.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3937)  
Tang,Y.T., Zhou,P. and Drmanac,R.T.  
AUTHORS Nucleic acids and polypeptides  
JOURNAL Patent: US 6569662-A 280 27-MAY-2003;  
FEATURES Location/Qualifiers  
source 1..3937  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 99.9%; Score 2911.4; DB 6; Length 3937;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACCTCGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGCTGTGTA 60  
DB 334 ATGGCCACCTCGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGCTGTGTA 393

QY 61 GGATCATTTGCTGTGTCTACAGAGCTGAGTCATTCACATGCTTTGGAAGTTGAAATC 120  
DB 394 GGATCATTTGCTGTGTCTACAGAGCTGAGTCATTCACATGCTTTGGAAGTTGAAATC 453

QY 121 AAAATGATGATGAAGAACGATGACAAAGCAGAAATGTTACAGAGTCCAAATGAG 180  
DB 454 AAAATGATGATGAAGAACGATGACAAAGCAGAAATGTTACAGAGTCCAAATGAG 513

QY 181 GTGAAATATCATTTGCCAATTTGAAACATCTTCTATCTTGAGACTTTTAACTTAATTTTGA 240  
DB 514 GTGAAATATCATTTGCCAATTTGAAACATCTTCTATCTTGAGACTTTTAACTTAATTTTGA 573

QY 241 GATAGCAATTAATGTATCTGTATTTGAATATGTCATTAATGAGAAATGAAAGATAT 300  
DB 574 GATAGCAATTAATGTATCTGTATTTGAATATGTCATTAATGAGAAATGAAAGATAT 633

QY 301 CTAAAGATAGAGTGAACCCCTTCACAAAATGAAAGTCCGACATTCATGACCAAGATC 360  
DB 634 CTAAAGATAGAGTGAACCCCTTCACAAAATGAAAGTCCGACATTCATGACCAAGATC 693

QY 361 ATCAGAGGATGTTGATCTTCATTCATGATGTTACTACACCGGAGCTTCACACTTCT 420  
DB 694 ATCAGAGGATGTTGATCTTCATTCATGATGTTACTACACCGGAGCTTCACACTTCT 753

QY 421 AACCTCCCTACTGACTCGTATATTTGAACATCAAGATTTGCTGATTTTGGCTGGCACTCA 480  
DB 754 AACCTCCCTACTGACTCGTATATTTGAACATCAAGATTTGCTGATTTTGGCTGGCACTCA 813

QY 481 CTGAATAATGCAATGAAGAAGCACTATACATTAATGTTGAACTCCCTAATCTAATTTCA 540  
DB 814 CTGAATAATGCAATGAAGAAGCACTATACATTAATGTTGAACTCCCTAATCTAATTTCA 873

QY 541 GAAATTTGCACTCGAAGTGCATGAGCTTGAATCTGAATGTTTGTGCTGGGCTGTATG 600  
DB 874 GAAATTTGCACTCGAAGTGCATGAGCTTGAATCTGAATGTTTGTGCTGGGCTGTATG 933

QY 601 TTTTATACATTAATGAGGAGCAACCTTCGACATGACATGACATGACATGACATGACAT 660  
DB 934 TTTTATACATTAATGAGGAGCAACCTTCGACATGACATGACATGACATGACATGACAT 993

QY 661 AATAAGTAGTATTTGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
DB 994 AATAAGTAGTATTTGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 1053

QY 721 CTTATTCACACAGTTACTTGTGAGAAATCCAGAGATCGTTTAAGTCTGCTCAGATTTG 780  
DB 1054 CTTATTCACACAGTTACTTGTGAGAAATCCAGAGATCGTTTAAGTCTGCTCAGATTTG 1113

QY 781 GACCATCTTTTATATGCCCCGGAATTTCTTCAACAAAAGATTAAGATTGAACTGTGAA 840  
DB 1114 GACCATCTTTTATATGCCCCGGAATTTCTTCAACAAAAGATTAAGATTGAACTGTGAA 1173

QY 841 GACTCAATGATAGTGGGATGCGCAATTTCTACTGCAATTAAGCTTCTTCCAGTATC 900  
DB 1174 GACTCAATGATAGTGGGATGCGCAATTTCTACTGCAATTAAGCTTCTTCCAGTATC 1233

QY 901 AGTATTAAGTGGTATTTTATTTGACAAAAGAACTTTGATTTGTCAGCCACTCCAAAT 960  
DB 1234 AGTATTAAGTGGTATTTTATTTGACAAAAGAACTTTGATTTGTCAGCCACTCCAAAT 1293

QY 961 AAAATGACTGATTTTCCAAAGAAATTAAGTCACTGATTTTCTTCTTCCAGAGATGGA 1020  
DB 1294 AAAATGACTGATTTTCCAAAGAAATTAAGTCACTGATTTTCTTCTTCCAGAGATGGA 1353

QY 1021 AACAGTTTATCTCACTGAGGAGAAATCAAGAAACAGTATAGTGAAGGGAAGATGA 1080  
DB 1354 AACAGTTTATCTCACTGAGGAGAAATCAAGAAACAGTATAGTGAAGGGAAGATGA 1413

QY 1081 ATTCAAGATGAGAAAGAAAGGCCATTTCTCGATATCTTGTGATAGCTTATCTCTGAT 1140  
DB 1414 ATTCAAGATGAGAAAGAAAGGCCATTTCTCGATATCTTGTGATAGCTTATCTCTGAT 1473

QY 1141 AGATCTGGCACTTCTAATAGTCAAGTCTCAAGCAAAAATATATCAATGAAAGATGAC 1200  
DB 1474 AGATCTGGCACTTCTAATAGTCAAGTCTCAAGCAAAAATATATCAATGAAAGATGAC 1353

QY 1201 TCACAGAAATGCTTCAAGTCTCAAGAAAGATCAAGAGAGAGTGAATGAAGAGGTAC 1260  
DB 1354 TCACAGAAATGCTTCAAGTCTCAAGAAAGATCAAGAGAGAGTGAATGAAGAGGTAC 1593

QY 1261 TCACCAAGACAAATGACCAATTTTAACTTTTAAAGAAAGACATCAAGATG 1320  
DB 1594 TCACCAAGACAAATGACCAATTTTAACTTTTAAAGAAAGACATCAAGATG 1653

QY 1321 TCTGATCTTTTGAAGACTGATTAACATCAAGCACTTCCAAATCATCTTGTCCAGGA 1380  
DB 1654 TCTGATCTTTTGAAGACTGATTAACATCAAGCACTTCCAAATCATCTTGTCCAGGA 1713

QY 1381 AAAAATCTCTTTCATTTGCAAGCCGACACCTCAAGCTGAAACGCTACAAAGTGT 1440  
DB 1714 AAAAATCTCTTTCATTTGCAAGCCGACACCTCAAGCTGAAACGCTACAAAGTGT 1773

QY 1441 GGGATCTGCAATTAATGCTCATTTAAGAAAACTACTGAATATGACAGATCAAGCCCA 1500  
DB 1774 GGGATCTGCAATTAATGCTCATTTAAGAAAACTACTGAATATGACAGATCAAGCCCA 1833

QY 1501 AACCGGGAATTCAGAGGCGCATCCAGATTTGCAAGAGGACAAATCAAAAATGCTGAGCT 1560  
DB 1834 AACCGGGAATTCAGAGGCGCATCCAGATTTGCAAGAGGACAAATCAAAAATGCTGAGCT 1893

QY 1561 GATACAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCACTTCTGTAAAACAGCA 1620  
DB 1894 GATACAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCACTTCTGTAAAACAGCA 1953

QY 1621 AATACCATGAATATATGATGCTGCACTTCAAGTAACTGAGATATCCAAAGAAATGT 1680  
DB 1954 AATACCATGAATATATGATGCTGCACTTCAAGTAACTGAGATATCCAAAGAAATGT 2013

QY 1681 GTTTTGGCTCAGATCTTCTTCTGAACAGAGCAAGATGAGGAGTATGAGGACACATAG 1740  
DB 2014 GTTTTGGCTCAGATCTTCTTCTGAACAGAGCAAGATGAGGAGTATGAGGACACATAG 2073

```

QY 1741 GGTATACAGATCGTACATTAAGAGCATTAACCTCCGTTGGTCTCAGAGATTAA 1800
DB 2074 GATTATCAGATCGTACATTAAGAGCATTAACCTCCGTTGGTCTCAGAGATTAA 2133
QY 1801 CCAATCAGACAGAAAAACAAAAAGGCTGTGTGAGCATTAAGATTCAAGAGAGGTGT 1860
DB 2134 CCAATCAGACAGAAAAACAAAAAGGCTGTGTGAGCATTAAGATTCAAGAGAGGTGT 2193
QY 1861 GTGAGAGCTGTGTAAAGAGATGATCTCAAGAAATATGTGAAAAGAGTCTTCAAGATATCT 1920
DB 2194 GTGAGAGCTGTGTAAAGAGATGATCTCAAGAAATATGTGAAAAGAGTCTTCAAGATATCT 2253
QY 1921 AGTGAATGAATATGATCACTATTTATTTATTCCAAAATGTGTAGAGTCTTCCCTGTGCT 1980
DB 2254 AGTGAATGAATATGATCACTATTTATTTATTCCAAAATGTGTAGAGTCTTCCCTGTGCT 2313
QY 1981 GATAAGCCACCTCTCAGCTACTGACACATCAGTAGGTAGAGCTTTGACAAATTTCCAGAA 2040
DB 2314 GATAAGCCACCTCTCAGCTACTGACACATCAGTAGGTAGAGCTTTGACAAATTTCCAGAA 2373
QY 2041 AATATCTGGCGAAATATCATATGTCTTCAGAGTGTGTACAGCTTTGTAAGATCTAATCT 2100
DB 2374 AATATCTGGCGAAATATCATATGTCTTCAGAGTGTGTACAGCTTTGTAAGATCTAATCT 2433
QY 2101 CCAAAATCACTTATTTTCAAGATATGCTAATATGCAATTTGATGAGAAATCTCCTGGT 2160
DB 2434 CCAAAATCACTTATTTTCAAGATATGCTAATATGCAATTTGATGAGAAATCTCCTGGT 2493
QY 2161 GCTGATTTTGAAGTGTGTGTATGATGAGGTTAAATATCAGAAAAACAGAGATTTCATT 2220
DB 2494 GCTGATTTTGAAGTGTGTGTATGATGAGGTTAAATATCAGAAAAACAGAGATTTCATT 2553
QY 2221 CAGGTGATTTGAAAAGACAGAGAACTTTACACTTTAAAAAGTGAAGTGAATTAATAC 2280
DB 2554 CAGGTGATTTGAAAAGACAGAGAACTTTACACTTTAAAAAGTGAAGTGAATTAATAC 2613
QY 2281 TTGAAAAGAGAGATTAATATGATATGACCAATGCTAATGAGGTCATGATATTTGTTA 2340
DB 2614 TTGAAAAGAGAGATTAATATGATATGACCAATGCTAATGAGGTCATGATATTTGTTA 2673
QY 2341 GCACTGGAATCCATATATTTCAAGAGAGAAAGAAAACTAGAGAGTCTCCCTTTTCCCA 2400
DB 2674 GCACTGGAATCCATATATTTCAAGAGAGAAAGAAAACTAGAGAGTCTCCCTTTTCCCA 2733
QY 2401 AATATCATATGAGAAAGAAACCTGTGTACTAGTTCACTTAAGGCTTATCACTCTCTCT 2460
DB 2734 AATATCATATGAGAAAGAAACCTGTGTACTAGTTCACTTAAGGCTTATCACTCTCTCT 2793
QY 2461 TCTGTGAGATTTCAAAATTAACCAAGAGAGATGAGCATCTTTCAAGAGATGAGCAT 2520
DB 2794 TCTGTGAGATTTCAAAATTAACCAAGAGAGATGAGCATCTTTCAAGAGATGAGCAT 2853
QY 2521 AGTGTGCTCTTCTCAACAGAGCAACCAATCTTAACTCCCTTATGTGTTCAAAATGAAGA 2580
DB 2854 AGTGTGCTCTTCTCAACAGAGCAACCAATCTTAACTCCCTTATGTGTTCAAAATGAAGA 2913
QY 2581 CTGTGTCTTCAACAGCTCTTGTGAACAGACATCTCTTATATAGTCTTAAAGATGT 2640
DB 2914 CTGTGTCTTCAACAGCTCTTGTGAACAGACATCTCTTATATAGTCTTAAAGATGT 2973
QY 2641 CTTCCTTAATCAGACAACTTTTGAATATCTTTTGTGAAAATATGTTGTGGGTGATA 2700
DB 2974 CTTCCTTAATCAGACAACTTTTGAATATCTTTTGTGAAAATATGTTGTGGGTGATA 3033
QY 2701 CAGTTAAGTGAAGAGCTGTGTGAGTTCAAGTTAATGATGAGTCCAGTTGTGTGAG 2760
DB 3034 CAGTTAAGTGAAGAGCTGTGTGAGTTCAAGTTAATGATGAGTCCAGTTGTGTGAG 3093
QY 2761 GAGAGAGTGTCTTATCAGTTATACCTCAACAAATGTGTAACAACTAAGATGAAGA 2820
DB 3094 GAGAGAGTGTCTTATCAGTTATACCTCAACAAATGTGTAACAACTAAGATGAAGA 3153
QY 2821 AATGAAAATTAACAGACTACATCAACAGAAATTAACAGTGTGTCTTCACTCTTTG 2880

```

```

DB 3154 AATGAAAATTAACAGACTACATCAACAGAAATTAACAGTGTGTCTTCACTCTTTG 3213
QY 2881 AATGTTTCTTAATCCGACTCTTAATTTTCAATTA 2913
DB 3214 AATGTTTCTTAATCCGACTCTTAATTTTCAATTA 3246

RESULT 3
HSTPKSAK
LOCUS HSTPKSAK 3092 bp mRNA linear PRI 27-MAY-1997
DEFINITION Homo sapiens mRNA for serine/threonine protein kinase SAK.
ACCESSION Y13115
VERSION Y13115.1 GI:2125813
KEYWORDS serine/threonine protein kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Karn, T., Holtrich, U., Wolf, G., Hock, B., Strubhardt, K. and
Kuebsamen-Waligmann, H.
TITLE Human SAK related to the PKC/polo family of cell cycle kinases
JOURNAL shows high mRNA expression in testis
AUTHORS Oncol. Rep. 4, 505-510 (1997)
TITLE 2 (bases 1 to 3092)
AUTHORS Karn, T.
REFERENCE Direct Submission
JOURNAL Submitted (09-MAY-1997) T. Karn, Chemotherapeutisches
Forschungsinstitut, Georg-Speyer-Haus, Paul-Ehrlich-Strasse 42-44,
D-60596 Frankfurt, FRG

FEATURES
source
1..3092
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="K3"
/tae="K3"
/tae_type="lung"
1..3092
/gene="SAK"
141..3053
/gene="SAK"
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="CAA73575.1"
/db_xref="GI:2125814"
/db_xref="GOA:O00444"
/db_xref="SPTREMBL:O00444"
/translation="MATCIQEKIEDPKFNGNLGKSPAGVYRAESIHNGLEVAIKMID
KKAMKACMOVORVKNVRIHQHOLKRPSTLELVNFEEDSNVYLVLEMCHNGMRYLK
NRYKPSERENARHPNHOITGMVLYHSHGILHRDITLSLLTRMWNITKIDRGALTO
LMPEKHYTLCTGPNVYISPEIATRSAGLESVDVSLGCMFTYLLIGRPFTDVKV
TLNKVLVDYEMPTLSIEAKDLIHQLRRNPADRLSSVLDHPMSNRSSTKSDI
GTVEDSIDSGATISTAITASSTSIISGLPEKRLILGQPLNPKMGSTGSPKSSDIF
SSSGGSPFYTOGNOETNSGGRGVIODAEPRPSRYLRAYSIDRSSTGSSOQAK
TYTMRCHSAEMLSVSKSGSENERERSYTDNNMNTNPFREKTSSTSGSEBERDNN
QHLNHLCEGKTPPPADPTPOTETVQAFGLQJNAHLKRTTETDISIPNDFQCHP
DLQKDTSKNAWDTIVKKNDSADNAHSVQKQNTKMTALSKVEIIQDVCPSDP
LSEQKTEGMEPPMPCYONKRLSITSPLVANHKEIRKOTKKAIVSIIIDSEVCVELV
KEYAOEYKVELVLOISDQNTITTYYPNGRGSPPLADPPSPDNIISRYSPNLBEKY
WRKYOASRFVOLVRSKSPKTYPIRYAKCLIMENSPGADPFVWFYDVKIHKTEDFI
QVIERKGSYTIKSSSEYNSLKEIKMPEDNABGHRICTALESTIISSEERTREAPF
PFIITGRKPGSTSSPSPVDSNTPTDRASPNRMVMSAASPTQAPILNBSMV
TNEGILGTTASGTDIISNSLKDCLPKSAQLKSVFVNKVGATQDTGSAVVOFNDG
SOLVVOAGVSSISYTSPNGQTTTRYGENEKLDPYIKOKLQCLSSILMFSNPBNH"

ORIGIN
Query Match 99.8%; Score 2906.6; DB 9; Length 3092;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2909; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AATGCACTGATCGAGGAGAAATCGAGATTTTAAAGTTGAATCTGCTGTGTA 60

```

```
Db 141 ATGGCGACTGTCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGTTGGTAA 200
QY 61 GGAATCTTGGTGGTGTCTTACAGAGTGAATCTTCACTGAGTTTGGAAATTTGCAATC 120
Db 201 GGAATCTTGGTGGTGTCTTACAGAGTGAATCTTCACTGAGTTTGGAAATTTGCAATC 260
QY 121 AAAATGATAGTAAAGAAAGCAATGTACAAAGCAGAAATGTACAGAGTCCAAATAG 180
Db 261 AAAATGATAGTAAAGAAAGCAATGTACAAAGCAGAAATGTACAGAGTCCAAATAG 320
QY 181 GTGAAATCATTTGGCAATTTGAAACATCTTTCTATCTTTGAGCTTTATTAATTTGAA 240
Db 321 GTGAAATCATTTGGCAATTTGAAACATCTTTCTATCTTTGAGCTTTATTAATTTGAA 380
QY 241 GATACCAATTAATGTGTATCTGTATTAGAAATGTSCATAATGTGAAATGAACAGTAT 300
Db 381 GATACCAATTAATGTGTATCTGTATTAGAAATGTSCATAATGTGAAATGAACAGTAT 440
QY 301 CTAAAGAAATGAGTAAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATC 360
Db 441 CTAAAGAAATGAGTAAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATC 500
QY 361 ATCAAGAGATGTGTATCTTCAATTCATGATTAATCAACCGGGACCTCAACTTTCT 420
Db 501 ATCAAGAGATGTGTATCTTCAATTCATGATTAATCAACCGGGACCTCAACTTTCT 560
QY 421 AACCTCTTACTGACTCGTAAATATGAAACATCAAGATTCGATTTGGGCTGGCACTCA 480
Db 561 AACCTCTTACTGACTCGTAAATATGAAACATCAAGATTCGATTTGGGCTGGCACTCA 620
QY 481 CTGAAATGCGACATGAAAGCACTATATAGTAACTGAACTCTTAATCTTCAATTTGCA 540
Db 621 CTGAAATGCGACATGAAAGCACTATATAGTAACTGAACTCTTAATCTTCAATTTGCA 680
QY 541 GAAATTCGCACTGCAAGTGCATGAGCTTTGAAATGTGATGTTTGGTCCCTGGCTGTATG 600
Db 681 GAAATTCGCACTGCAAGTGCATGAGCTTTGAAATGTGATGTTTGGTCCCTGGCTGTATG 740
QY 601 TTTTATCATTTACTTATCGGGAAGCAACCTTGCACATGCACTGACAGTCAAGAACATTA 660
Db 741 TTTTATCATTTACTTATCGGGAAGCAACCTTGCACATGCACTGACAGTCAAGAACATTA 800
QY 661 AATTAAGTATGATTTGCAAGATTAAGAAATGCCATCTTTTGTGCAATAGAGCCAAAGAC 720
Db 801 AATTAAGTATGATTTGCAAGATTAAGAAATGCCATCTTTTGTGCAATAGAGCCAAAGAC 860
QY 721 CTATTCACACAGTTACTTGTAAGAAATCCAGCAGATCGTTTAACTGTCTTCAATATG 780
Db 861 CTATTCACACAGTTACTTGTAAGAAATCCAGCAGATCGTTTAACTGTCTTCAATATG 920
QY 781 GACCATCTTTTATGTCGCCAAATCTTTCACAAAGAAAGTAAAGTTTGAAGAACTGTGAA 840
Db 921 GACCATCTTTTATGTCGCCAAATCTTTCACAAAGAAAGTAAAGTTTGAAGAACTGTGAA 980
QY 841 GACTCAATGATAGTGGGAGATGCCCAATTTTACTGCAATTAACGCTTCTTCCAGTACC 900
Db 981 GACTCAATGATAGTGGGAGATGCCCAATTTTACTGCAATTAACGCTTCTTCCAGTACC 1040
QY 901 AGTATAGTGTAGTATTTTGAACAAAGAAAGACTTTGATTGTGAGGCCCTCCCAAT 960
Db 1041 AGTATAGTGTAGTATTTTGAACAAAGAAAGACTTTGATTGTGAGGCCCTCCCAAT 1100
QY 961 AAAATGACGTATTTCCAAAGAAATGAAGTTCAACTGATTTTCTTCCAGAGATGGA 1020
Db 1101 AAAATGACGTATTTCCAAAGAAATGAAGTTCAACTGATTTTCTTCCAGAGATGGA 1160
QY 1021 AACAGTTTATTAATCACTGAGTGGGAAATCAAGAAACAGTAAATAGTGAAGGGAGATGA 1080
Db 1161 AACAGTTTATTAATCACTGAGTGGGAAATCAAGAAACAGTAAATAGTGAAGGGAGATGA 1220
QY 1081 ATTCAAGATGCAAGAAAGGCCCACTTTCGATACCTTGTAGAGCTTATTCCTGAT 1140
Db 1221 ATTCAAGATGCAAGAAAGGCCCACTTTCGATACCTTGTAGAGCTTATTCCTGAT 1280
QY 1141 AGATCTGGGACCTTCTAATATGTCAAGTCTCAAGCAAAACATATACATGGAACGATGTAC 1200
Db 1281 AGATCTGGGACCTTCTAATATGTCAAGTCTCAAGCAAAACATATACATGGAACGATGTAC 1340
QY 1201 TCAGCAAAATGCTTTCACTGTCTCAAAAGATCAGAGAGAGTGAATGAAGAGGTAC 1260
Db 1341 TCAGCAAAATGCTTTCACTGTCTCAAAAGATCAGAGAGAGTGAATGAAGAGGTAC 1400
QY 1261 TCACCCACAGACAAATGCCCAACTTTTAACTTCTTTAAAGAAAGACATCCAGTAT 1320
Db 1401 TCACCCACAGACAAATGCCCAACTTTTAACTTCTTTAAAGAAAGACATCCAGTAT 1460
QY 1321 TCTGATCTTTTGAAGAGCTGATTAACATCAAGCACTCTCAATCATCTTTGTCCAGA 1380
Db 1461 TCTGATCTTTTGAAGAGCTGATTAACATCAAGCACTCTCAATCATCTTTGTCCAGA 1520
QY 1381 AAAATCTCTTTTCCATTTTGGAGACCTCGACACTCGACATGAAACCGTACACAGTGT 1440
Db 1521 AAAATCTCTTTTCCATTTTGGAGACCTCGACACTCGACATGAAACCGTACACAGTGT 1580
QY 1441 GGAATCTGCAAAATTAATGCTCATTTAAGAAACCTGTAATATGACAGATACGCCA 1500
Db 1581 GGAATCTGCAAAATTAATGCTCATTTAAGAAACCTGTAATATGACAGATACGCCA 1640
QY 1501 AACCGGACCTTCCAGGCGCATCCAGATTTGCAAGAGACATCAAAATAATGCTGGACT 1560
Db 1641 AACCGGACCTTCCAGGCGCATCCAGATTTGCAAGAGACATCAAAATAATGCTGGACT 1700
QY 1561 GATACAAATGCAAAAGAAAGTCTGATGCTTGTGATTAAGCAATCTGTTAAACAGCA 1620
Db 1701 GATACAAATGCAAAAGAAAGTCTGATGCTTGTGATTAAGCAATCTGTTAAACAGCA 1760
QY 1621 AATACATGAAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATG 1680
Db 1761 AATACATGAAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATG 1820
QY 1681 GTTTTGGCTCAAGTCCCTCTTCTGAAACAGAGCAAGATGAGGGTATGAGCCACCATG 1740
Db 1821 GTTTTGGCTCAAGTCCCTCTTCTGAAACAGAGCAAGATGAGGGTATGAGCCACCATG 1880
QY 1741 GGTATATCAAGATGATTAAGAAACATTAATCTCCGTTGGTGTCTCAAGGTTAA 1800
Db 1881 GGTATATCAAGATGATTAAGAAACATTAATCTCCGTTGGTGTCTCAAGGTTAA 1940
QY 1801 CCAATCAGACAGAAACCAAAAGGCTGTGAGTACATCTTGATTCAGAGAGGTGTGT 1860
Db 1941 CCAATCAGACAGAAACCAAAAGGCTGTGAGTACATCTTGATTCAGAGAGGTGTGT 2000
QY 1861 GTGAGCTGTGAAGAGATGATCTCAAGAAATGTGAAGAAAGTCTTCAAGTATCT 1920
Db 2001 GTGAGCTGTGAAGAGATGATCTCAAGAAATGTGAAGAAAGTCTTCAAGTATCT 2060
QY 1921 AGTATGAGAAATACATCTATTTATTAATCCAAATGCTGTAAGAGTTTCTCTTGTCT 1980
Db 2061 AGTATGAGAAATACATCTATTTATTAATCCAAATGCTGTAAGAGTTTCTCTTGTCT 2120
QY 1981 GATPAGCAACCTTCACTACTGACAAATCAGTAGTACAGTTTGAATTAACCGAA 2040
Db 2121 GATPAGCAACCTTCACTACTGACAAATCAGTAGTACAGTTTGAATTAACCGAA 2180
QY 2041 AAATCTGGCGGAAATATCAATATGCTTCAAGTTTGTACAGCTTGAAGATCTAAATCT 2100
Db 2181 AAATCTGGCGGAAATATCAATATGCTTCAAGTTTGTACAGCTTGAAGATCTAAATCT 2240
QY 2101 CCCAAATCACTTATTTTAAAGATATGCTTAAATGCAATTTGATGAGAAATCTCTGCT 2160
Db 2241 CCCAAATCACTTATTTTAAAGATATGCTTAAATGCAATTTGATGAGAAATCTCTGCT 2300
QY 2161 GGTGATTTTGAAGTTTGTATGATGAGGTAAATAACAAAGACAGAAATTCATT 2220
Db 2301 GGTGATTTTGAAGTTTGTATGATGAGGTAAATAACAAAGACAGAAATTCATT 2360
```

Oy 2221 CAGGTGATGAAAGACAGGAGAGCTTCACTTTTAAAAAGTGAAGTGAATTAATAC 2280  
 Db 2361 CAGGTGATGAAAGACAGGAGAGCTTCACTTTTAAAAAGTGAAGTGAATTAATAC 2420  
 Oy 2281 TTGAAAAGAGAGTAAATATGATATGACCATGCTAATGAGGCTCATGTAATTTGTTA 2340  
 Db 2421 TTGAAAAGAGAGTAAATATGATATGACCATGCTAATGAGGCTCATGTAATTTGTTA 2480  
 Oy 2341 GGAATGGAATTCATATTTTCAAGAGAGGAAAGAAATTAAGAGAGTCCCTTTTCCCA 2400  
 Db 2481 GGAATGGAATTCATATTTTCAAGAGAGGAAAGAAATTAAGAGAGTCCCTTTTCCCA 2540  
 Oy 2401 ATATATATAGAGAAAGAAAGCTGCTAGTACTAGTCTAAGGCTTATCACTCTCTCT 2460  
 Db 2541 ATATATATAGAGAAAGAAAGCTGCTAGTACTAGTCTAAGGCTTATCACTCTCTCT 2600  
 Oy 2461 TCTGTGATTCATATTTTCAAGAGAGTATAGAGATCTTTTCAAGAGATGCTATGAT 2520  
 Db 2601 TCTGTGATTCATATTTTCAAGAGAGTATAGAGATCTTTTCAAGAGATGCTATGAT 2660  
 Oy 2521 AATGCTGCTTCTCCCAACAGAGGACCAATCCCTTATCCCTTATGCTTACAAATGAAGA 2580  
 Db 2661 AATGCTGCTTCTCCCAACAGAGGACCAATCCCTTATCCCTTATGCTTACAAATGAAGA 2720  
 Oy 2581 CTGCTGCTTACAACTACAGCTTCTGGAACAGATCTCTTCTAATAGTCTTAAAGATTGT 2640  
 Db 2721 CTGCTGCTTACAACTACAGCTTCTGGAACAGATCTCTTCTAATAGTCTTAAAGATTGT 2780  
 Oy 2641 CTGCTGCTTACAACTACAGCTTCTGGAACAGATCTCTTCTAATAGTCTTAAAGATTGT 2700  
 Db 2781 CTGCTGCTTACAACTACAGCTTCTGGAACAGATCTCTTCTAATAGTCTTAAAGATTGT 2840  
 Oy 2701 CAGTTAATAGAGAGAGTCTGAGGCTTCACTTTTAAAGAGGCTCCAGATGCTGAGGAG 2760  
 Db 2841 CAGTTAATAGAGAGAGTCTGAGGCTTCACTTTTAAAGAGGCTCCAGATGCTGAGGAG 2900  
 Oy 2761 GGAAGAGTCTTCTATCTATCTATATACCTCAACAAATGCTCAACATAGAGTGAAGA 2820  
 Db 2901 GGAAGAGTCTTCTATCTATCTATATACCTCAACAAATGCTCAACATAGAGTGAAGA 2960  
 Oy 2821 AATGAAAATTAACAGAGTATCAATCAAAAGAAATTAAGAGTCTGCTTCTTCACTTTTG 2880  
 Db 2961 AATGAAAATTAACAGAGTATCAATCAAAAGAAATTAAGAGTCTGCTTCTTCACTTTTG 3020  
 Oy 2881 ATGTTTCTTAATCCGACTCTTAATTTTCAATGA 2913  
 Db 3021 ATGTTTCTTAATCCGACTCTTAATTTTCAATGA 3053

RESULT 4  
 AB006972 3331 bp mRNA linear PRI 18-OCT-2001  
 LOCUS AB006972 Homo sapiens mRNA for Sak, complete cds.  
 DEFINITION AB006972  
 ACCSSION AB006972.1 GI:16215695  
 VERSION Sak.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Yamaehita, Y., Kajigaya, S., Yoshida, K., Ueno, S., Ota, J., Omine, K.,  
 Ueda, M., Miyazato, A., Ohya, K., Kitemura, T., Ozawa, K. and Mano, H.  
 Sak serine-threonine kinase acts as an effector of Tec tyrosine  
 kinase  
 J. Biol. Chem. 276 (42), 39012-39020 (2001)  
 MEDLINE 21466437  
 PUBMED 11489907  
 REFERENCE 2 (bases 1 to 3331)  
 AUTHORS Mano, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-SEP-1997) Hiroyuki Mano, Jichi Medical School,  
 Department of Molecular Biology, 331-1 Yakushiji,

FEATURES  
 source  
 Location/Qualifiers  
 1..3331  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="K562"  
 /tissue\_type="blood"  
 189..3101  
 /codon\_start=1  
 /product="Sak"  
 /protein\_id="BAB69958.1"  
 /db\_xref="GI:16215696"  
 /translation="MATCGECIEDPKVGNLKGKSPAGVPAABSHGLFVALIMID  
 KRAMYKAGVQVONBVKIHCOLKHPISLIELNYFEDSNVTVLVBMCGENMYLK  
 NKVKEPSENEBAHFHMOIITMLYLHSHILHRDLTSLMLLRNNIKIADPLATVO  
 LKMPHKVTLGCTPNYISPEIATRSAGLESDVMSLCGMYTLLIGRPFDVTVN  
 TLNKVVLADYEMPTFISIEADLIHOLLRNPADRLSLSSVDHPMSNSSTKSDL  
 GTVEDSIDSGHATISTATTAASSTSISSGLPKRRLTIGOLPINKMTVPKMSSTDF  
 SSSGDSNFPYTOGMOETNSGSRGAVIODAERPHSRIRBARVSSDRGTSNRQSAK  
 TYTWERCHSAEMLSVSKSGSENERKISPTDNNAITENFKEKTSSSSGSERPDNN  
 QALSNHLCPGKTPFPADPTPQETVQWFGNLQINHLRKTEDYDISPNDFQSH  
 DLQKDTSKNAATDTLVKONSIDSNASVQONTKMYTALSKKEIIOQECVFGSDP  
 LSEOSKTRGMEPMDGJOURTLASITSPVLAHLKPIROKTKAAVVISLSEVCEVLY  
 KEVASOEYKEVLISSDGNITTIYPNGRGEPPLADRPSPDINSRVSFNDLPKXY  
 WKQVYASFPVOLLASKSPKITTYFTRVAKCIEMNSRGADPWPVMDYKIKTRDPI  
 QVIEKTKSYTLKSSSEVNSLKEEIKMTDHPHNEGRITLAEESIISSEKRTKRAFP  
 FPIIIGRKPGSTSSPKALSPSPVDSNPTRDRASPNRNVMSDASDPAPLNPBMV  
 TNEGGLTTTASGTDISSNSLDCPKSKHOLKSVFVKVNGVATOLTSAGVWVQFNDG  
 SOLVVGAVSSISYTSPPNGQTRVGENEKLPIYIKQKLCISILMFSPNPFH"

ORIGIN  
 Query Match 99.7%; Score 2905; DB 9; Length 3331;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2908; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Oy 1 ATGGCGACCTGCATGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGCTGTGATA 60  
 Db 189 ATGGCGACCTGCATGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGCTGTGATA 248  
 Oy 61 GGATCATTTTGTGCTGTCTACAGAGCTGAGTCCATTCACACTGCTTGTGAAAGTTGAAATC 120  
 Db 249 GGATCATTTTGTGCTGTCTACAGAGCTGAGTCCATTCACAGAGTTTGAAGTTGAAATC 308  
 Oy 121 AAAATGATGATPAAGAAAGCCATGTACAAAGAGAGATGTACAGAGTCCAAAATGAG 180  
 Db 309 AAAATGATGATPAAGAAAGCCATGTACAAAGAGAGATGTACAGAGTCCAAAATGAG 368  
 Oy 181 GTGAAAATATCATTTGCCAATTTGAAACATCTCTATCTTGAGACTTATATATTTGAA 240  
 Db 369 GTGAAAATATCATTTGCCAATTTGAAACATCTCTATCTTGAGACTTATATATTTGAA 428  
 Oy 241 GATAGCAATTAATGTATCTGATTAATGAATATGTCATATAGAGAAATGAACAGATAT 300  
 Db 429 GATAGCAATTAATGTATCTGATTAATGAATATGTCATATAGAGAAATGAACAGATAT 488  
 Oy 301 CTAAAGATATGATGAAAGCCCTTCTCGAATAATGAAGTTCGACACTTCATGCAAGATC 360  
 Db 489 CTAAAGATATGATGAAAGCCCTTCTCGAATAATGAAGTTCGACACTTCATGCAAGATC 548  
 Oy 361 ATCAACAGGAGATTTGATCTTCTCATGTGATATCTACACCGGAGCTCAACACTTCT 420  
 Db 549 ATCAACAGGAGATTTGATCTTCTCATGTGATATCTACACCGGAGCTCAACACTTCT 608  
 Oy 421 AACCTCTACTGACTCTGATATATGAACATCAAGATTTGCTGATTTGGCTGGCAACTCAA 480  
 Db 609 AACCTCTACTGACTCTGATATATGAACATCAAGATTTGCTGATTTGGCTGGCAACTCAA 668  
 Oy 481 CTGAAAATCCCATGAAAGACATATACATTAATGTGAAAGTCTCTAATCAATTTGACA 540  
 Db 669 CTGAAAATCCCATGAAAGACATATACATTAATGTGAAAGTCTCTAATCAATTTGACA 728

```
QY 541 GAAATGGCACTCGAAGTGCACATGGCCTTGAATCTGATGTTGGTCCCTGGGCTGTATG 600
DB 729 GAAATGGCACTCGAAGTGCACATGGCCTTGAATCTGATGTTGGTCCCTGGGCTGTATG 788
QY 601 TTTTATACATTACTTATCGGAGAACACCCCTTGCACATGACAGATCAAGAACATTA 660
DB 789 TTTTATACATTACTTATCGGAGAACACCCCTTGCACATGACAGATCAAGAACATTA 848
QY 661 AATAAAGTATGATGGAGATTAAGAAATGCACTTTTGTCAATATAGAGGCCAAGAC 720
DB 849 AATAAAGTATGATGGAGATTAAGAAATGCACTTTTGTCAATATAGAGGCCAAGAC 908
QY 721 CTATTCACACGATTACTTCGTAGAAATCGACAGATGTTAAGTCTGTCTTCAGTATG 780
DB 909 CTATTCACACGATTACTTCGTAGAAATCGACAGATGTTAAGTCTGTCTTCAGTATG 968
QY 781 GACCATCTTTTATGATCCCAAAATTTCTCACAAAAAGTAAGAATTTAGAACTGTGAA 840
DB 969 GACCATCTTTTATGATCCCAAAATTTCTCACAAAAAGTAAGAATTTAGAACTGTGAA 1028
QY 841 GACTCAATGATAGTGGGCATGCGCAAAATTTCTACATGCAATTAAGCTTTTCCAGTACC 900
DB 1029 GACTCAATGATAGTGGGCATGCGCAAAATTTCTACATGCAATTAAGCTTTTCCAGTACC 1088
QY 901 AGTATAGTGTAGTATTTTGAACAAAAGAACCTTTGATGATGACCACTCCCAAT 960
DB 1089 AGTATAGTGTAGTATTTTGAACAAAAGAACCTTTGATGATGACCACTCCCAAT 1148
QY 961 AAATATGCTATATTTCCAAAGAAATTAAGTTCAATGATTTTCTTCCAGAGATGGA 1020
DB 1149 AAATATGCTATATTTCCAAAGAAATTAAGTTCAATGATTTTCTTCCAGAGATGGA 1208
QY 1021 AACAATTTTATATCTAGTGGGGAATCAAGAAACGATATATGTAAGGGGAGAGTA 1080
DB 1209 AACAATTTTATATCTAGTGGGGAATCAAGAAACGATATATGTAAGGGGAGAGTA 1268
QY 1081 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTGAGCTTATTTCTCTGAT 1140
DB 1269 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTGAGCTTATTTCTCTGAT 1328
QY 1141 AGATCTGGCACTTCTATATAGTCACTCAAGCAAAAACATATATCAAGATGTAAC 1200
DB 1329 AGATCTGGCACTTCTATATAGTCACTCAAGCAAAAACATATATCAAGATGTAAC 1388
QY 1201 TCAGCAGAATGCTTCACTGTCCAAAAGATCAAGAGAGGTGAATCAAGAGAGTAC 1260
DB 1389 TCAGCAGAATGCTTCACTGTCCAAAAGATCAAGAGAGGTGAATCAAGAGAGTAC 1448
QY 1261 TCACCAACAGACAAATGSCAAATTTTAACTTCTTAAGAAAGAACATCCAGTATG 1320
DB 1449 TCACCAACAGACAAATGSCAAATTTTAACTTCTTAAGAAAGAACATCCAGTATG 1508
QY 1321 TTTGATCTTTTGAAGAACCTGTATTAACAATCAAGCACTTCCATCATCTTTGTCCAGA 1380
DB 1509 TCTGATCTTTTGAAGAACCTGTATTAACAATCAAGCACTTCCATCATCTTTGTCCAGA 1568
QY 1381 AAAATCTCTTTTCAATTTGAGAACCCGACACTCGACTGAACCGTATCAACAGTGT 1440
DB 1569 AAAATCTCTTTTCAATTTGAGAACCCGACACTCGACTGAACCGTATCAACAGTGT 1628
QY 1441 GGGAACTGCAAAATTAATGCTATTTAAGAAAACTACTGAATATGACAGATGAGCCA 1500
DB 1629 GGGAACTGCAAAATTAATGCTATTTAAGAAAACTACTGAATATGACAGATGAGCCA 1668
QY 1501 AACCGGAGCTTCCAGGGGCATTCAGATTTGCAAGAGACATCAAAAAATGCTGTGACT 1560
DB 1689 AACCGGAGCTTCCAGGGGCATTCAGATTTGCAAGAGACATCAAAAAATGCTGTGACT 1748
QY 1561 GATACAAAAAGTCAAAAAAGACCTGATGCTCTGATATATGCACTTCTGTATTAACAGCA 1620
DB 1749 GATACAAAAAGTCAAAAAAGACCTGATGCTCTGATATATGCACTTCTGTATTAACAGCA 1808

QY 1621 AATACCATGAATATATATGACTGCACTTCAAGTAAACCTGAGATATCCACAAGATGT 1680
DB 1809 AATACCATGAATATATATGACTGCACTTCAAGTAAACCTGAGATATCCACAAGATGT 1868
QY 1681 GTTTTGGCTGCAATCTCTTTCTGACAGAGCAAGACTAGGGGTATGAGCCACATG 1740
DB 1869 GTTTTGGCTGCAATCTCTTTCTGACAGAGCAAGACTAGGGGTATGAGCCACATG 1928
QY 1741 GGTATCAGAAATCGTACATTAAGAAAGCATTAATCCCGTGGTGTCTCACAGGTTAAA 1800
DB 1929 GGTATCAGAAATCGTACATTAAGAAAGCATTAATCCCGTGGTGTCTCACAGGTTAAA 1988
QY 1801 CCAATCAGCAGAAAAACAAAAGGCTGTGTGAGCATATTTGATTCAGAGAGGTGT 1860
DB 1989 CCAATCAGCAGAAAAACAAAAGGCTGTGTGAGCATATTTGATTCAGAGAGGTGT 2048
QY 1861 GTGAGCTTGTAAAGAGATGCACTCAAGAAATATGGAAGAAAGTTCTTCAATATCT 1920
DB 2049 GTGAGCTTGTAAAGAGATGCACTCAAGAAATATGGAAGAAAGTTCTTCAATATCT 2108
QY 1921 AGTATGAAATATCGATCACTATTTATATCAATGGGTAGAGTTTCTCTTGTCT 1980
DB 2109 AGTATGAAATATCGATCACTATTTATATCAATGGGTAGAGTTTCTCTTGTCT 2168
QY 1981 GATAGACACCCCTCACCTTGAACAATCATGTAAGTACAGCTTTGACAAATTTACAGAA 2040
DB 2169 GATAGACACCCCTCACCTTGAACAATCATGTAAGTACAGCTTTGACAAATTTACAGAA 2228
QY 2041 AAATCTGGCAAAATATCAATATGCTTCCAGTTTGAACAGCTTGAATCAATATCT 2100
DB 2229 AAATCTGGCAAAATATCAATATGCTTCCAGTTTGAACAGCTTGAATCAATATCT 2288
QY 2101 CCCAAATCACTTATTTTACAGAAATATGATGCAATTTGATGAGAAATTCCTGTGT 2160
DB 2289 CCCAAATCACTTATTTTACAGAAATATGATGCAATTTGATGAGAAATTCCTGTGT 2348
QY 2161 GCTGATTTGAGGTTGTGTTTATGATGGGTATAAATATACAAAACGAAGATTTCA 2220
DB 2349 GCTGATTTGAGGTTGTGTTTATGATGGGTATAAATATACAAAACGAAGATTTCA 2408
QY 2221 CAGGTATTTGAAAGACAGGGAAGTCTTACATTTAAAGTAAAGTGAATTAATAGC 2280
DB 2409 CAGGTATTTGAAAGACAGGGAAGTCTTACATTTAAAGTAAAGTGAATTAATAGC 2468
QY 2281 TTGAAAGAGAGATTAATAATGTATATGACACATGCTATATAGGGTCACTGATTTGTTA 2340
DB 2469 TTGAAAGAGAGATTAATAATGTATATGACACATGCTATATAGGGTCACTGATTTGTTA 2528
QY 2341 GCACGTGAATCAATATTTTCAAGAGAGAAAGAAACTAGAGAGTCTCCCTTTCCCA 2400
DB 2529 GCACGTGAATCAATATTTTCAAGAGAGAAAGAAACTAGAGAGTCTCCCTTTCCCA 2588
QY 2401 ATATATCATAGGAAGAAACCTGTATATTAAGTTCACTTAAGGCTTATCACTCTCTCT 2460
DB 2589 ATATATCATAGGAAGAAACCTGTGTATTAAGTTCACTTAAGGCTTATCACTCTCTCT 2648
QY 2461 TCTGTGATTTCAATTAACCAACGAGATGAGCACTTTCAACGAATGTCAATGAT 2520
DB 2649 TCTGTGATTTCAATTAACCAACGAGATGAGCACTTTCAACGAATGTCAATGAT 2708
QY 2521 AGTGTGCTTTCTCAACACAGGCAACCAATCTTAAATCCCTTATAGGTTTCAATGAGGA 2580
DB 2709 AGTGTGCTTTCTCAACACAGGCAACCAATCTTAAATCCCTTATAGGTTTCAATGAGGA 2768
QY 2581 CTGTGCTTAAACAATCAAGCTTCTGAAACAGACATCTCTTAAATAGTCTAAAGATGT 2640
DB 2769 CTGTGCTTAAACAATCAAGCTTCTGAAACAGACATCTCTTAAATAGTCTAAAGATGT 2828
QY 2641 CTTCCTAAATCAGCAAACTTTTGAATCTGTTTGTGAAAAATGTTGTTGGGCTACA 2700
DB 2829 CTTCCTAAATCAGCAAACTTTTGAATCTGTTTGTGAAAAATGTTGTTGGGCTACA 2888
QY 2701 CAGTTAACTAGTGAAGCTGTGTGGGTTCAATTAATGATGGGTCCAGTTGTGTGAG 2760
```

|||||  
Db 2889 CAGTTAACTAGTGGAGCTGTGGGTTCACTTTAAAGATGGGCTCCAGTTGGTGTGCGAG 2948  
Qy 2761 GCAAGAGTGTCTTCTATCACTTATACCTCAACCAATGTCTCAACAACTGATGTGAGAA 2820  
Db 2949 GCAAGGTGTCTTCTATCACTTATACCTCAACCAATGTCTCAACAACTGATGTGAGAA 3008  
Qy 2821 AATGAAATTTACGACCTACATCAACGAATTAACAGTGTCTGCTTCACTCTTTTGG 2880  
Db 3009 AATGAAATTTACGACCTACATCAACGAATTAACAGTGTCTGCTTCACTCTTTTGG 3068  
Qy 2881 ATGTTTCTAATCCGACTCTCTAATTTTCATTTGA 2913  
Db 3069 ATGTTTCTAATCCGACTCTCTAATTTTCATTTGA 3101  
RESULT 5  
AR083569 3447 bp DNA 1linear PAT 01-SEP-2000  
LOCUS Sequence 3 from patent US 5976893.  
ACCESSION AR083569  
VERSION AR083569.1 GI:10010342  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3447)  
AUTHORS Dennis,J.W., Heffernan,M. and Fode,C.  
TITLE Methods for identifying binding partners, agonists, and antagonist  
JOURNAL Patent: US 5976893-A 3 02-NOV-1999;  
FEATURES  
source 1..3447  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 64.5%; Score 1879; DB 6; Length 3447;  
Best Local Similarity 80.9%; Pred.No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

Qy 1 ATGGGCACTGTCATCGGGAGAGAAATCGAGATTTTAAAGTTGGAATCTGCTTGTA 60  
Db 206 ATGGGCGCTGTCATCGGGAGAGAAATCGAGATTTTAAAGTTGGAATCTGCTTGTA 265  
Qy 61 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACACCTGTTTGAAGTTGCAATC 120  
Db 266 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACACCTGTTTGAAGTTGCAATC 325  
Qy 121 AAAATGATGATTAAGAAAGCCATGTAAGAGAGAAATGTTACAGAGATCCAAATGAG 180  
Db 326 AAAATGATGATTAAGAAAGCCATGTAAGAGAGAAATGTTACAGAGATCCAAATGAG 385  
Qy 181 GTGAAATATCATTTGCAATTTGAAACATCTTCTATCTTGAGCTTTATATCTTTGAA 240  
Db 386 GTGAAATATCATTTGCAATTTGAAACATCTTCTTGGAGCTTATATTTACTTTGAA 445  
Qy 241 GATAGCAATTAATGTATCTGTGATTAAGAAATGTGCAATGAGAAATGAAACAGTAT 300  
Db 446 GATAGCAATTAATGTATCTGTGATTAAGAAATGTGCAATGAGAAATGAAACAGTAT 505  
Qy 301 CTAAAGAAATGATGAAACCTTTCTGAGAAATGAGAGCTGACACTTTCATGACCGAGATC 360  
Db 506 CTAAAGAAATGATGAAACCTTTCTGAGAAATGAGAGCTGACACTTTCATGACCGAGAT 565  
Qy 361 ATCAAGAGATGTTGATCTTCACTTCTCATGTATCTACACCGGAGCTCACTTTCT 420  
Db 566 ATCAAGAGATGTTGATCTTCACTTCTCATGTATCTACACCGGAGCTCACTTTCT 625  
Qy 421 AACCTCTACTGATCTGTAATATGAAATCAATCAAGATTTGAGCTGCGCACTCAA 480  
Db 626 AACCTCTACTGATCTGTAATATGAAATCAATCAAGATTTGAGCTGCGCACTCAA 685

Qy 481 CTGAAATATCCATATGAAAGCACTATACATTAATGTGAATCTTAATCACTTTCACCA 540  
Db 686 TTGAATATCCATATGAAAGCACTATACATCTGTGGGACCTCTTAATTAATTTTACCA 745  
Qy 541 GAAATTCGACTGAAAGTGCACATGGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600  
Db 746 GAAATTCGAACTGAAAGTGCACATGGAATCTGATGTTTGGTCCCTGGGCTGTATG 805  
Qy 601 TTTTATACATTAATTAATGGGAGACACCTTCGACATCGACAGAGTCAAGAACATTA 660  
Db 806 TCTTATACATTAATTAATGGGAGACACCTTTTGAACATCGACAGTCAAGAACATTA 865  
Qy 661 AATTAAGTATGTTGGCAGATTAATGAAATGCCATCTTTTGTTCATATAGAGCCAGAGAC 720  
Db 866 AATTAAGTATGTTGGCAGATTAATGAAATGCCATCTTTTGTTCATATAGAGCCAGAGAC 925  
Qy 721 CTATATCACAGTTAATCTGTAGAAATCCAGAGATCGTTTAAGTGTCTTCAATTTG 780  
Db 926 CTATATCACAGTTAATCTGTAGAAATCCGTGAGATCGGTTAAGTGTCTTCTGTGTTG 985  
Qy 781 GACATCTCTTTATATGTCGGAAATCTTCAAGAAATGAAATTAAGTATAGAGCTGTGAA 840  
Db 986 GACATCTCTTTATATGTCGGAAATCTTCAAGAAATGAAATTAAGTATAGAGCTGTGAA 1045  
Qy 841 GACTCAATGATAGTGGGATGCCAAATTTCTACTGCAATTAAGCTTCTTCAAGTACC 900  
Db 1046 GACTCAATGATAGTGGGATGCCAAATTTCTACTGCAATTAAGCTTCTTCTGTATAC 1105  
Qy 901 AGTATAGTGTAGTTTATTTTGAAGAAAGACATTTTGTATGTTGTCAGCTCCCAAT 960  
Db 1106 AGTATAGTGTAGTTTATTTTGAAGAAAGACATTTTGTATGTTGTCAGCTCCCAAT 1162  
Qy 961 AAAATGACCTGTATTTCCAAAGAAATGAAATGTTTCTTCTTCAAGAGATGGA 1020  
Db 1163 AAAATGACCTGTATTTCCAAAGAAATGAAATGTTTCTTCTTCAAGAGATGGA 1219  
Qy 1021 AACAGTTTATATCACTAGTGGGAAAT-----CAAGAAACAGTATATGTAAGGGGA 1074  
Db 1220 AACAGTTTATATCACTAGTGGGAAATCCAGAAACAGTATATGTAAGGGGACGGGG 1279  
Qy 1075 AAGATTAATTAAGTGCAGAAAGGACCAATTCATCTGATCTTCTGATAGCTTATTC 1134  
Db 1280 AAGATTAATTAAGTGCAGAAAGGACCGCATTCATCTGATCTGATAGCTTATTC 1339  
Qy 1135 TCTGATAGTCTGCACTTCTAATAGTCACTTCAAGCAAAACATATCAATGGAACGA 1194  
Db 1340 TCTGATAGTCTGCACTTCTAATAGTCACTTCAAGCAAAACATATCAATGGAACGA 1396  
Qy 1195 TGTCACTGAGCAAAATGCTTCAAGTGTCAAAAGATCAGAGAGGTGAAATGAAGAG 1254  
Db 1397 TGTCACTGAGCAAAATGCTTCAAGGCTTAAGATC----- 1434  
Qy 1255 AGGTACTCACCAAGACAAATGCAATTTTAACTTTTAAAGAAAGACATCC 1314  
Db 1435 ----- 1434  
Qy 1315 AGTATGTTGATCTTTGAAAGACCTGATTAACATCAAGCACTCTGCAATCATCTTTGT 1374  
Db 1435 -----ACTGATGAAATTCACACAGTTCCATCATCATCTTGT 1471  
Qy 1375 CCAGGAAAAATCTCTTTTCACTTTCAGAGCCGAGACCTCAGAGTGAACCGTACACAG 1434  
Db 1472 CTAGAAAAATCTCTTTTCACTTTCAGAGCCGAGACCTCAGAGTGAATGATAGCAG 1531  
Qy 1435 TGGTTTGGGAATCTGCAATTAATGCTCTTTTAAAGAAAACTAATGAAATATGACAGATC 1494  
Db 1532 TGGTTTGGGAATCTGCAATTAATGCTCTTTTAAAGAAAACTAATGAAATATGACAGATC 1591  
Qy 1495 AGCCCAAAACCGGAGCTTCCAGGCGCATCCAGATTTTGGAGAAAGACATCAAAAATGCGC 1554  
Db 1592 AGCCCAAAACAGAGATTTCCAGGAGCTATCCAGATTTTGC--AGGACAGTTTGAAGAACGCT 1648  
Qy 1555 TGAAGTATCAAAAAGTCAAAAAGTCAAACTGTGATGCTTGTGATATGACATTTCTGTAAAA 1614

1649 TGGAGTGCAGAGAGCCAGCAAGATCTGATCTTCTGCGCAATGTTCAAGCTGTAAAG 1708  
1645 CAGCAAAATACAGAAATATATGATCTGCACTTCAAGTAAACCTGAGATTAATCAACAA 1674  
1709 CAGCTGAGTGCAGAAATATATGATCTGCACTTCAAGTAAACCTGAGTATCCACAG 1768  
1675 GAATGTGTTTTTGGCTCAGATCTCTTCTGTAAGAGAGCAAGATAGAGGGATGAGACCA 1734  
1769 GA-----GCCGGGCTATCATCTCATTTCTGAAGAGCAAGATAGAGATGAGATCG 1822  
1735 CCATGGGGTTTTCAGATCTGATCTTAAAGCACTTAATCTCGTTGGTGTCTCAGAG 1794  
1823 ACCTGGGTTTACAGAAACCTACCTTAAGAAATTAATCTCTCTGATGTGCTCAGAGA 1882  
1795 TTAAGCAATCAGACAGAAACCAAAAGGCTGTGTGAGCACTACTGTTTCAAGAGAG 1854  
1883 TTAAGCAATCAGACAGAAACCAAAAGGCTGTGTGAGCACTACTGTTTCAAGAGAG 1942  
1855 GTGTGTGTGAGCTTGTAAAGAGATGATCTCAAGATATGTGAAGAAATCTTCAAG 1914  
1943 GTGTGTGTGAGCTTGTGAAGAGATGTGCTGTGAAGATATGTGAAGAAATGCTTCAAG 2002  
1915 ATATCTGATGAGAAATAGATCACTATTTATTTATCAAAATGCTGTGAGAGGTTTCT 1974  
2003 ATATCTGATGAGAGATGATCACTATTTATTTATCAAAATGCTGTGAGAGGTTTCT 2062  
1975 GTGTGTGTGAGCACTCTCACTACTGACATGATGATGATGATGATGATGATGATGAT 2034  
2063 GTGTGTGTGAGCACT 2122  
2035 CCGAAGAAATATCTGAGGAGAAATATCAATATGCTTCAAGGTTGTGATGATGATGAT 2094  
2123 CCGAAGAAATATCTGAGGAGAAATATCAATATGCTTCAAGGTTGTGATGATGATGAT 2182  
2095 AAATCTCCCAAAATCACTTATTTTAAAGATATGATGATGATGATGATGATGATGAT 2154  
2183 AAATCTCCCAAAATCACTTATTTTAAAGATATGATGATGATGATGATGATGATGAT 2242  
2155 CCGAGTGTGATTTTGAAGGTTGTGATGATGATGATGATGATGATGATGATGATGAT 2214  
2243 CCGAGTGTGATTTTGAAGGTTGTGATGATGATGATGATGATGATGATGATGATGAT 2302  
2215 TTCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2274  
2303 TTAATTCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2362  
2275 AATAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2334  
2363 ACCAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2422  
2335 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394  
2423 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482  
2395 TTCCCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2454  
2483 TTCCCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2542  
2455 CCTCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514  
2543 CCTCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2599  
2515 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
2600 GTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659  
2575 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634  
2660 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2708  
2635 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694

Db 2709 ----GTCTCTTAATCTGCACAGCTTTTGAATCTGTTTTGTGAAAAATGTGGTTGG 2764  
Qy 2695 GCTACACAGTTAATCTAGTGAAGCTGTGTGGTTCAAGTTTAAAGAGGGCCAGTTGGTT 2754  
Db 2765 GCTACACAGTTAATCTAGTGAAGCTGTGTGGTTCAAGTTTAAAGAGGGCCAGTTGGTT 2824  
Qy 2755 GTGACAGCGAGAGTGTCTTCTATCAAGTTTAACTTCAACCAATGTCTCAACCACTAGTAT 2814  
Db 2825 GTTCAGAGCGAGAGTATCTTCTCAATCAGTTTACATCAACAGATGTCTCAGCAACTAGTAT 2884  
Qy 2815 GGAGAAATAGAAATATACAGATCAATCAACAGAAATTAACAGTGTCTGTCTTCCATC 2874  
Db 2885 GGAGAAATAGAAATATACAGATCAATCAACAGAAATTAACAGTGTCTGTCTTCCATC 2944  
Qy 2875 CTTTGAATGTTTCTTAATCTCAATCTCAATTTTCAATTTGA 2913  
Db 2945 CTTTGAATGTTTCTTAATCTCAATCTCAATTTTGA 2983

RESULT 6  
156871 LOCUS 156871 3447 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 3 from patent US 5650501.  
ACCESSION 156871  
VERSION 156871.1 GI:2477284  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3447)  
AUTHORS Dennis,J.W., Heffernan,M. and Fode,C.  
TITLE Serine/threonine kinase and nucleic acids encoding same  
JOURNAL Patent: US 5650501-A 3 22-JUL-1997;  
FEATURES  
source location/Qualifiers  
1..3447  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 64.5%; Score 1879; DB 6; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;  
Qy 1 ATGCGACCTGCAATCGGGAGAGATGAGAGATTTTAAAGTGAATCTGTGGTAA 60  
Db 206 ATGCGGCGTGCATCGGGAGAGATGAGAGATTTTAAAGTGAATCTGTGGTAA 265  
Qy 61 GGATCATTTGTGTGTCTTCAAGAGCTGATGATCACTGTTTGAAGTTGCAATC 120  
Db 266 GGATCATTTGTGTGTCTTCAAGAGCTGATGATCACTGTTTGAAGTTGCAATC 325  
Qy 121 AAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
Db 326 AAATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385  
Qy 181 GTGAAAATATCATTTGCAATTTGAACATCTTCTATCTTGAAGCTTTTAACTTTTGA 240  
Db 386 GTGAAAATATCATTTGCAAGTTGAACATCTTCTATCTTGAAGCTTTTAACTTTTGA 445  
Qy 241 GATAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Db 446 GATAGCAATATATCTTCACTGTGTGATGATGATGATGATGATGATGATGATGATGAT 505  
Qy 301 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 506 CTGAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565  
Qy 361 ATACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 566 ATACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 625  
Qy 421 AACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

Db 626 AACATCTTACTTACGCGGAATATGAACATVAAAAATTTGTCGACTTGGACTAGCAAGCGAC 685  
 QY 481 CTGAAAAATCCACATGAAGACCTATATATTTGGAATCTCCCTAATCAATTTTCCACA 540  
 Db 686 TTGAATATCCACATGAAGACCTATATATTTGGAATCTCCCTAATCAATTTTCCACA 745  
 QY 541 GAAATTCGACATGAGTGCACATGAGCTTGAATCTGATGTTGATCCCTGGGCTGTATG 600  
 Db 746 GAAATTCGACATGAGTGCACATGAGCTTGAATCTGATGTTGATCCCTGGGCTGTATG 805  
 QY 601 TTTTATACATTAATTATGCGGAGACCACTTCTGACATGACATGATCAAGAACATTA 660  
 Db 806 TCTTATACCTTACTTATGGAAGACCACTTCTGACATGACATGATCAAGAACATTA 865  
 QY 661 AATTAAGTATGATGACATTAATGAAGCCATCTTTTGTGATTAAGGCGAAGAC 720  
 Db 866 AACAAAGTATGATGACATTAATGAAGCCATCTTTTGTGATTAAGGCGAAGAC 925  
 QY 721 CTTAATCACCAATTAATGATGAGAAATTCAGACATGATGATTAAGTCTGCTTCAATTTG 780  
 Db 926 CTATTCACACATTAATGATGAGAAATTCAGACATGATGATTAAGTCTGCTTCAATTTG 985  
 QY 781 GACATCTCTTATATGTCGCGAAATTCCTCAACAAAAGTAAAGATTAAGGAACTGTGAA 840  
 Db 986 GACATCTCTTATATGTCGCGAAATTCCTCAACAAAAGTAAAGATTAAGGAACTGTGAA 1045  
 QY 841 GACTCAATGATATGAGGACATGACAAATTTCTACTGCAATTAACATCTTCCATATAC 900  
 Db 1046 GACTCAATGATATGAGGACATGACAAATTTCTACTGCAATTAACATCTTCCATATAC 1105  
 QY 901 AGTATTAAGTATGATTAATTTGCAAAAGAAAGCTTTGATGATGAGGACATCCCAAT 960  
 Db 1106 AGTATTAAGTATGATTAATTTGCAAAAGAAAGCTTTGATGATGAGGACATCCCAAT 1162  
 QY 961 AAAATGACTGATTTTCCAAAGAAATTAAGTCAATGATTTTCTTCAAGAGATGA 1020  
 Db 1163 AAAATGACTGATTTTCCAAAGAAATTAAGTCAATGATTTTCTTCAAGAGATGA 1219  
 QY 1021 AACAGTTTATATCTCATGAGGGAAT-----CAAGAAACCATTAATGAGAGGGA 1074  
 Db 1220 AATTAATTTTGTATCTCAATGAGGGAATTCAGAAACAAAGATTAATGAGAGGGA 1279  
 QY 1075 AAGATTAATCAAGATSCAGAAAGGACCAATTCGATACCTTGTAGGCTTATTC 1134  
 Db 1280 AAGATTAATCAAGATSCAGAAAGGACCAATTCGATACCTTGTAGGCTTATTC 1339  
 QY 1135 TCTGATAGATCTGGCACTTCTATATAGTCAATCAAGCAAAAATATTAATGAGAGCA 1194  
 Db 1340 TCTGATAGATCTGGCACTTCTATATAGTCAATCAAGCAAAAATATTAATGAGAGCA 1396  
 QY 1195 TGTCACTCAAGCAAAATGCTTTCAGTGTCCAAAAGATCAAGAGAGGTGAAAATGAAGAG 1254  
 Db 1397 TGTCACTCAAGCAAAATGCTTTCAGTGTCCAAAAGATCAAGAGAGGTGAAAATGAAGAG 1434  
 QY 1255 AAGTACTCAACCAAGACAAATGCAATTTTAACTTTTAAAGAAAAGACATCC 1314  
 Db 1435 ----- 1434  
 QY 1315 AGTATGTTCTGATCTTTTGAAGACCTGATTAACATCAAGCACTTCAATCATCTTTGT 1374  
 Db 1435 -----ACTGATTAAGAAATCAACACAGTTCCATATCATTTGT 1471  
 QY 1375 CCAAGAAAAATCTCTTTTCAATTTGCAAGCCGACATCTCAAGTAAACCTGTAACAAG 1434  
 Db 1472 CTAGAAAAATCTCTTTTCAATTTGCAAGCCGACATCTCAAGTAAATGTAAGACAG 1531  
 QY 1435 TGGTTTGGGAATCTGCAAAATTAATGCTCAATTAAGAAAACTAGATATGACAGATC 1494  
 Db 1532 TGGTTTGGGAATCTGCAAAATTAATGCTCAATTAAGAAAACTAGATATGACAGATC 1591  
 QY 1495 AGCCCAAAACCGGAGACTTCAAGGCGCATCAAGATTTTCAAGAGACATCAAAAATGCGC 1554  
 Db 1592 AGCCCAAAACAGAGATTTTCAAGGACTATCAAGATTTGC---AGGACAGGTTACGAAAGCGCT 1648

QY 1555 TGGACTGATACAAAGTCAAAAAGAACTGATGCTTCTGATATGACATTTCTGTAATA 1614  
 Db 1649 TGGACTGATACAAAGTCAAAAAGAACTGATGCTTCTGATATGACATTTCTGTAATA 1708  
 QY 1615 CAGCAAAATACATGAAATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1674  
 Db 1709 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1768  
 QY 1675 GAAATGATTTTGGCTCAGATCTCTTCTTGAACAGAGAAAGATGAGGATGAGGCA 1734  
 Db 1769 GA-----GCCGAGGCTCAGATCTCTTCTTGAACAGAGAAAGATGAGGATGAGGCA 1822  
 QY 1735 CCATGGGCTTATCAGAAATGATCAATTAAGAGACATTAACCTCCGTTGATGATCAGAG 1794  
 Db 1823 ACATGAGGCTTATCAGAAATGATCAATTAAGAGACATTAACCTCCGTTGATGATCAGAG 1882  
 QY 1795 TTAACCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCATTAATTAAGAGAG 1854  
 Db 1883 TTAACCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCATTAATTAAGAGAG 1942  
 QY 1855 GTGTGTGTGAGCTTGTATAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1914  
 Db 1943 GTGTGTGTGAGCTTGTATAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2002  
 QY 1915 ATATCTAGATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1974  
 Db 2003 ATATCTAGATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2062  
 QY 1975 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2034  
 Db 2063 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2122  
 QY 2035 CCAAGAAAAATACCTGCGGAAATATCAATATGCTTCAAGTGTGATGATGATGATGATGAT 2094  
 Db 2123 CCAAGAAAAATACCTGCGGAAATATCAATATGCTTCAAGTGTGATGATGATGATGATGAT 2182  
 QY 2095 AAATCTCCCAAAATCACTATTTTCAAGATATGATGATGATGATGATGATGATGATGATGAT 2154  
 Db 2183 AAATCTCCCAAAATCACTATTTTCAAGATATGATGATGATGATGATGATGATGATGATGAT 2242  
 QY 2155 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2214  
 Db 2243 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2302  
 QY 2215 TTCAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2274  
 Db 2303 TTCAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2362  
 QY 2275 AATAGCTGAAAGAGAGATTAATAATGATATGATGATGATGATGATGATGATGATGATGAT 2334  
 Db 2363 AATAGCTGAAAGAGAGATTAATAATGATATGATGATGATGATGATGATGATGATGATGAT 2422  
 QY 2335 TGTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394  
 Db 2423 TGTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482  
 QY 2395 TTCCCAATATCATAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGATGATGAT 2454  
 Db 2483 TTCCCAATATCATAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGATGATGAT 2542  
 QY 2455 CCTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514  
 Db 2543 CCTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2599  
 QY 2515 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
 Db 2600 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659  
 QY 2575 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634  
 Db 2660 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2708

Query 2635 GATTGCTCTTCCAAATCAGACAACCTTTGAAATCTGTTTGTGAAAAAATGTTGGTGG 2634  
Db 2709 -----GTCTTCTTAATCTGCACAGCTTTTGAAATCTGTTTGTGAAAAAATGTTGGTGG 2764  
Qy 2685 GCTTACACAGTTAATAGTAGAGCTGTGTGGGTTCAGTTAATGATGGGTCCAGTGGTT 2754  
Db 2765 GCTTACACAGTTAATAGTAGAGCTGTGTGGGTTCAGTTAATGATGGGTTCAGTGGTT 2824  
Qy 2755 GTGACGAGGAGAGTGTCTTCTATCAGTTAATCCCTACCAAAATGTCACAACTAGGTAT 2814  
Db 2825 GTCCAGGACGAGATCTTCTCATGATTAACATCCACAGATGGTCCAGAACTAGGTAT 2884  
Qy 2815 GAGAAAAATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTGTCCATC 2874  
Db 2885 GAGAAAAATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTGTCCATC 2944  
Qy 2875 CTTTGTGATGTTTCTTATCCGACTCTCTAATTTTCATGGA 2913  
Db 2945 CTTTGTGATGTTTCTTATCCGACTCTCTAATTTTCATGGA 2983

RESULT 7  
AX306014 3447 bp DNA linear PAT 11-DEC-2001  
LOCUS Sequence 765 from Patent WO018188.  
DEFINITION AX306014  
ACCESSION AX306014 GI:17645355  
VERSION AX306014.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Bukayocia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T., and Ishii, Y.  
AUTHORS Method for examining ischemic conditions  
TITLE Patent: WO 018188-A 765 22-NOV-2001;  
JOURNAL School Juridical Person Nihon University (JP)  
FEATURES  
source location/Qualifiers  
1..3447  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Query Match 64.5%; Score 1879; DB 6; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

Db 566 ATCAGAGAAATGTATATCTTCAATTCATGGCAATTCGACCGGGACCTCACCTCTT 625  
Qy 421 AACCTCTCTAGCTCTGTAAATGAAATCAGATGTTGCTGATTTTGGGTGGCACTCA 480  
Db 626 AACATTTTACTTACGGGAAATGAAATGAAATTTGCTGATTTTGGGTGGCACTCA 685  
Qy 481 CTGAAATGCAATGAAATGCAATTAATATATGTAATGTAATGTAATGTAATGTAATG 540  
Db 686 TTGAATATGCAATGAAATGCAATTAATATGTAATGTAATGTAATGTAATGTAATG 745  
Qy 541 GAAATGCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 600  
Db 746 GAAATGCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 805  
Qy 601 TTTTATACATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 660  
Db 806 TCTTATACATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 865  
Qy 661 AATTAAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 720  
Db 866 AACAAAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 925  
Qy 721 CTTATTCACAGTTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 780  
Db 926 CTTATTCACAGTTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 985  
Qy 781 GACCACTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 840  
Db 986 GACCACTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1045  
Qy 841 GACCACTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 900  
Db 1046 GACCACTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1105  
Qy 901 AGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 960  
Db 1106 AGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1162  
Qy 961 AAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1020  
Db 1163 AAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1219  
Qy 1021 AACATTTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1074  
Db 1220 AACATTTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1279  
Qy 1075 AAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1134  
Db 1280 AAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1339  
Qy 1135 TCTGATGATCTGCACTTCTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1194  
Db 1340 TCTGATGATCTGCACTTCTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1396  
Qy 1195 TCTGATGATCTGCACTTCTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1254  
Db 1397 TCTGATGATCTGCACTTCTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1434  
Qy 1255 AAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1314  
Db 1435 AAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1434  
Qy 1315 AGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1374  
Db 1435 AGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1471  
Qy 1375 CCGGAAAAATCTCTTTCATTTTGAAGCCGACACTGTAATGTAATGTAATGTAATGTAATGTA 1434  
Db 1472 CCGGAAAAATCTCTTTCATTTTGAAGCCGACACTGTAATGTAATGTAATGTAATGTAATGTAATG 1531  
Qy 1435 TGGTTGGGAATCTGCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1494

Db 1532 TGGTTGGGAATCTGCAAAATGAATGCTCATTTAGAGAAACTAATGACCAACCGTT 1591  
 Qy 1455 AGCCCAACCGGGAATTCAGAGGCCATCCAGATTTGGAGAGAGACATCAAAAATATGCC 1554  
 Db 1592 AGCCCAACAGAGATTTCCAGAGCTATCCAGATTTGCG---AGGACCGTTACGAAACGCT 1648  
 Qy 1555 TGGACTGATACAAAGTCAAAAGAACTGTGATGCTTGTGATATGCAATTCGTGAAA 1614  
 Db 1649 TGGATGACACAGAGCCAGAGAAATGCTGATCTTCTGCCAAATGCTCATGCTGTAAAG 1708  
 Qy 1615 CAGCAAAATACAGATATATATGACTGCACTTCAAGTAACTGAGATTAATCCAA 1674  
 Db 1709 CAGCTGAGTCCATGAAATATACAGATGCAATACATAGCTGAGGCTGATGCGACAG 1768  
 Qy 1675 GAATGTGTTTTGGCTCAGATCTCTTTTGTGAACAGAGAACTAGGGGTATGAGCCN 1734  
 Db 1769 GA-----GCCGGGGCTCATCTCTCAATCTGAAACAAAGCAAAATAGAAATGAGTGG 1822  
 Qy 1735 CATAGGGGTTATCAGATGCTGATATTAAGAGCAATTAATCTCCGTGGTGTGACAGG 1794  
 Db 1823 ACATGGGTTACAGAAACCTTACAGAAATTAATCAATCTCTCTGATGCTGACAG 1882  
 Qy 1795 TTAACCAATACAGACAGAAACCAAAAGGCTGTGTGAGCAATCTTGATTCAGAGAG 1854  
 Db 1883 TTAAGCCCAATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1942  
 Qy 1855 GTGTGTGTGAGCTTGTAAAGAGTATGATCTCAAGATATGTGAAGAATGTTCTTCAG 1914  
 Db 1943 GTGTGTGTGAGCTTGTGAGAGGTGTGTGAGATATGTGAAGAATGTTCTTCAG 2002  
 Qy 1915 ATATGATGATGAGAAATACGATCACTATTTATTAATCCAAATGTTGAGAGTTCCT 1974  
 Db 2003 ATATGATGATGAGAAATACGATCACTATTTATTAATCCCAAGATGAGAGAGCTTCCT 2062  
 Qy 1975 CTGTGTGTGAGCAACCTCACTCACTGCAACATCAGATGAGTACAGCTTTGACATTTA 2034  
 Db 2063 CTGTGTGTGAGCAACCTCCCTGCTGCTACTGCAACATCAGATGAGTACAGCTTTGACATTTA 2122  
 Qy 2035 CCAGAAATATCTGGCGAAATATCAATATGCTTCCAGGTTGTACAGCTTTGATAGATTT 2094  
 Db 2123 CCAGAAATATCTGGCGAAATATCAATATGCTTCCAGATTCATTCAGCTTACATGATCT 2182  
 Qy 2095 AAATCTCCAAATACCTATTTTCAAGATATGCTAAATGCAATTTGATGAGAGATCT 2154  
 Db 2183 AAAATCTCCAAATACCTATTTTCAAGATATGCTAAATGCAATTTGATGAGAGATTTCT 2242  
 Qy 2155 CTTGTGTGTGATTTTGTAGGTTGTTTATGATGAGGTTAAATATCACAAACAGAGAT 2214  
 Db 2243 CTTGTGTGTGATTTTGTAGGTTGTTTATGATGAGGTTAAATATCATAAATCTGAAAT 2302  
 Qy 2215 TTCAATCAGGTGATGAGAAAGACAGGGAAGCTTACACTTTAAAGATGAAAGTGAAGTT 2274  
 Db 2303 TTAATTCACATTAATGAGAAACAGGGAATCTTAAATTTAAAGATGAAATGATGAGTT 2362  
 Qy 2275 AATACCTGAAAGAGAGAGATTAAGATATGATGAGCAATGCTAATGAGGTCATGATAT 2334  
 Db 2363 ACCACCTGAAAGAGAGAGATTAAGATATGATGAGCAATGCTAATGAGGTCATGATAT 2422  
 Qy 2335 TGTTTAGCACTGGAATCCATTAATTTCAAGAGAGAAAGAAATAGAGTCTCCCTTT 2394  
 Db 2423 TGTCTGTGATCTGGAATCTGTATCTGTAGAGAGAAAGAGAGAGAGGCTTCTTCAATTC 2482  
 Qy 2395 TTCCCAATATATAGAGAGAAAGAAAGCTGTGATGATGATGATGATGATGATGATGAT 2454  
 Db 2483 TTCCCAATATATATAGAGAGAAAGAAAGCTGTGATGATGATGATGATGATGATGAT 2542  
 Qy 2455 CTTCTGTGTGATTTCAAAATTAACCAAGAGAGATGAGATCTTTCAAGAAATGATGAT 2514  
 Db 2543 CTTCTGTGTGATTTCAAAATTAACCAAGAGAGATGAGATCTTTCAAGAAATGATGAT 2599  
 Qy 2515 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
 Db 2600 GTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659

Qy 2575 GAAGACCTTGTGCTTACAACTACAGCTTCTGGAACAGACATCTCTTATAGTCTAATA 2634  
 Db 2660 GAAGACCTTGTGCTTACAACTACAGCTTCTGGAACAGACATCTCTTATAGTCTAATA 2708  
 Qy 2635 GATGCTTCTTCTTAAATCAGACCAACTTTGAAATCTGTTTTGTGAAAAATATGTTGGTGG 2694  
 Db 2709 ----GTTCTCTTAAATCTGACACAGCTTTGAAATCTGTTTTGTGAAAAATATGTTGGTGG 2764  
 Qy 2695 GGTACACAGTTAACTAGTGTGAGCTGTGTGCTTCAATTAATGATGAGGCTCCAGTTGGTT 2754  
 Db 2765 GGTACACAGCTTAACTAGTGTGAGCTGTGTGCTTCAATTAATGATGAGGCTCCAGTTGGTT 2824  
 Qy 2755 GTGACAGAGAGAGTCTTCTTATACAGTTATACCTCAACCAATATGCTCAACATAGGTT 2814  
 Db 2825 GTGACAGAGAGAGTCTTCTTATACAGTTATACCTCAACCAATATGCTCAACATAGGTT 2884  
 Qy 2815 GGAGAAATGAAATTAATCAGACATCAATCAAAAGAAATTAACAGTGTCTGTTCATC 2874  
 Db 2885 GGAGAAATGAAATTAATCAGACATCAATCAAAAGAAATTAACAGTGTCTGTTCATC 2944  
 Qy 2875 CTTTGTGATGTTTTCTTAATCCAGCTCTTAATTTCTTGA 2913  
 Db 2945 CTTCTGATGTTTTCTTAATCCAGCTCTTAATTTCTTGA 2983

RESULT 8  
 MUSKAKA 3447 bp mRNA linear ROD 24-JUN-1994  
 LOCUS MUSKAKA 3447 bp mRNA linear ROD 24-JUN-1994  
 DEFINITION Mus musculus serine/threonine kinase (sak-a) mRNA, complete cds.  
 ACCESSION L29479  
 VERSION L29479.1 GI:487869  
 KEYWORDS protein kinase; serine/threonine kinase.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3447)  
 FODE, C., MOTTO, B., YOUSEFI, S., HEFFERNAN, M. and DENNIS, J. W.  
 Sak, a murine protein-kinase/threonine kinase that is related to the Drosophila polo kinase and involved in cell proliferation  
 Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6388-6392 (1994)  
 JOURNAL MEDLINE 94294387  
 PUBMED 8022793  
 COMMENT Original source text: Mus musculus (strain DBA/2) Lymphoid tumor CDNA to mRNA.  
 FEATURES  
 source location/Qualifiers  
 1..3447 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="DBA/2"  
 /db\_xref="taxon:10090"  
 /cell\_line="D3JW25"  
 /cell\_type="Transformed cell"  
 /tissue\_type="Lymphoid tumor"  
 1..3447 /gene="sak-a"  
 1..205 /gene="sak-a"  
 206..2983 /gene="sak-a"  
 /note="sak-a and sak-b are alternate splices with divergent COOH ends beginning at nucleotide 1454; putative"  
 /product="serine/threonine kinase"  
 /protein\_id="AAC37648.1"  
 /db\_xref="GI:487870"  
 /translation="MACTGRIEDPKGNLGGKSPAGYRAHSIHGLSVATIMID KKAATKAGVORVONVEKIHCOLKHPVLELYNPEDNNVYLVLEKCHNGEKRYLK NRMKPFSEARHFMHQIITGMLYHSGILHRDLTSLNILLTRMNKIKADFGATG LNMPEKVTLCGTPNYISPIATRSAGLSIDISLSCMSVTLIGRPDPDVTGN TLNKVVLADYEMPATLSREADLTHQLLRBPADRLTSSVLDHPFMSRNPSPKSDV

GTVEDMSDGSHTLSTTTTASGCTSLSGSLIDRLRLVGPPLNFKITVEQKNNSDPS  
 SGGNSNCTOMGNPBOBANSRGRVTEIEAERPHRYLRASHSDRASPQSRKAT  
 YSVYRCHSVSMLEKPRSLIDENSHNHCLGKTPFPADQTPQKMEYQOFGNLOMY  
 AHGETNHEHTVSPNRDQDYPDLOTLNNAWTDTRASNADTASNVHVAQOLSMKY  
 MSANHHKPEVMPQEPGLHPHSEQSKNRSSESTLQSKPLBSITSPILAHLLKPTROK  
 TKAAVVS1LDSEVCEVLEBECASBGVYEVLOISDGMLETVVYPNDRGPPLDRP  
 PLPTDMSRYSFNPLEKTYRWKQOYASRPTLOLRSRKTPTTFTYACCTIMENSPGA  
 DEEVMFYDAKIHKTENLIIHIEKTOISYLNKRENEVTSLEKEVVMYDHAENGRIC  
 LSLSEVISEEKRSGRSFPPRIIVGRKRNSTSPALVAPVDPSCQBOASRLS  
 VNSAARPPQSPGLSPSTTVBEGHATKATRGVSSSLPKSNOLKSFVKNVGMATQ  
 LNSGATWVPVNPNGSOVLVOAGVSSISYSPDQOTTRKGENEKLPEYIKOKJCSSLIL  
 LKPSNFTENFQ  
 2384..3447  
 /gene="sak-a"

3'UTR

ORIGIN

Query Match 64.5%; Score 1879; DB 10; Length 3447;  
 Best Local Similarity 80.9%; Pred. No. 0;  
 Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

QY 1 ATGGCGACCTGTCATCGGGAGAGAGATCGAGATTTTAAAGTTGGAAATTCGCTGTGTAA 60  
 DB 206 ATGGCGGGCGTGCATCGGGAGAGAGATCGAGATTTTAAAGTTGGAAATTCGCTGTGTAA 265  
 QY 61 GGATCATTTGCTGGTGTCTTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 120  
 DB 266 GGATCATTTGCTGGTGTCTTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 325  
 QY 121 AAAATGATGATTAAGAAAGCATGTACAAAGCAGAAATGGTACAGAGATGCCAAATGAG 180  
 DB 326 AAAATGATGATTAAGAAAGCATGTACAAAGCTGGAATGGTACAGAGATGCCAAATGAG 385  
 QY 181 GTGAAAATACATTCGCAATTTGAAACATTCCTTCTATCTTTGAGCTTTTAACTATTTTGA 240  
 DB 386 GTGAAAATACATTCGCAATTTGAAACATTCCTTCTTGTGAGCTCTATATTTACTTTGA 445  
 QY 241 GATAGCAATTATGTGTATCTGTATTTAGAAATGTGCCATTAATGAAATTAAGACAGTAT 300  
 DB 446 GATAGCAATTATGTGTATCTGTATTTAGAAATGTGCCATTAATGAAATTAAGACAGTAT 505  
 QY 301 CTAAAGAAATGAGAGAAACCCCTTCTCAGAAATTAAGAGCTGACACTTCATGACACGATC 360  
 DB 506 CTGAAAGAAAGAAATGAAGCCCTTCTCAGAAAGGAGAGCTAGGCACTTGAATGACAGAT 565  
 QY 361 ATCAGAGGATGTGTATCTTCTCATTCATGATTAATCTACACGGAGACTTCACACTTCT 420  
 DB 566 ATCAGAGGATGTGTATCTTCTCATTCATGATTAATCTACACGGAGACTTCACACTTCT 625  
 QY 421 AACTCTCTACGCTGTATTAATGAACATCAAGATTGCTGATTTGGGCTGGCACTGAA 480  
 DB 626 AACTCTCTACGCTGTATTAATGAACATCAAGATTGCTGATTTGGGCTGGCACTGAA 685  
 QY 481 CTGAAATATGCATGAAAGCACTATACATTAATGTGAACCTCTTAACTCATCTTGTACCA 540  
 DB 686 TTGAATATGCATGAAAGCACTATTAATGTGAACCTCTTAACTCATCTTGTGTACCA 745  
 QY 541 GAAATTTGCCATCGAGAGTGCACATGAGCCCTTGAATCTGAATGTTGGTCCCTGGGCTGATG 600  
 DB 746 GAAATTTGCCATCGAGAGTGCACATGAGCTTGAATCTGAATGTTGGTCCCTGGGCTGATG 805  
 QY 601 TTTTATACATTAATGAGGAGACACACCTTTCACACTGACACAGTCAAGACACATTA 660  
 DB 806 TCTTATAGCTTACTTATGGAAGACACCTTTTGAACATGACACAGTCAAGACACATTA 865  
 QY 661 AATAAAGTATGTTGGCAGATTATGAATGCCATCTTTTGTCAATAGAGGCCAAGAG 720  
 DB 866 AACAAGAGTATCTGGCAGATTATGAATGCAAGCTTTTGTCAAGAGGCCAAGAG 925  
 QY 721 CTATATTCACAGTTACTTGTGAATTCACAGATGATTGTTAAAGTGTCTTCAAGATTTG 780  
 DB 926 CTATATTCACAGTTACTTGTGAATTCACAGATGATTGTTAAAGTGTCTTCAAGATTTG 985  
 QY 781 GACCATCTTTTATGTCCGAATTTCTTCAACAAAAGTAAAGTTTAAAGAACTGTGGA 840

DB 986 GACCATCTTTTATGTCCGAATTCCTTCAACAAAGATTAAGACGTAGGACTGTAGAG 1045  
 QY 841 GACTCAATTGATAGTGGCATGCGACATTTTCACTGCAAAATTAACGCTTTCGATCC 900  
 DB 1046 GACTCAATTGATAGTGGCATGCGACATTTTCACTGCAAAATTAACGCTTTCGATCC 1105  
 QY 901 AGTAAAGTGTATGTTTATTTGACAAAGAAAGACTTTTGTATGTCAGCCACTCCAAAT 960  
 DB 1106 AGTTAAGTGTGAGCCCTTACTTAC--AGAAAGCTTTGTGTGTCAACACTTCCAAAT 1162  
 QY 961 AAAATGATGATTTTCAAAAGATTAAGTCAACTGATTTTCTTTCAGAGATGGA 1020  
 DB 1163 AAAATGATGATTTTCAAAAGATTAAGTCAACTGATTTTCTTTCAGAGATGGA 1219  
 QY 1021 AACAGTTTATATCTCAGTGGGAAAT-----CAAGAAACAGTAAATAGTGAAGGGA 1074  
 DB 1220 AATTAATTTTGTATCAATAGGGAATTCAGAAACAAAGACTAATATAGGGAAGGGG 1279  
 QY 1075 AGAGTAAATTCAGATGAGAAAGAAAGCCACATTCCTGATACCTTCTGATGCTTATTC 1134  
 DB 1280 AGAGTAAATTCAGATGAGAAAGAAAGCCACATTCCTGATACCTTCTGATGCTTATTC 1339  
 QY 1135 TCTGATAGTCTGCACTTCTAATAGTCAAGTCTCAAGCAAAACATATACATGGAACGA 1194  
 DB 1340 TCTGATAGTCTGCAAGCCCTCTAA---TCAAGTCTCAAGCAAAACATATCTGATGGAACGT 1396  
 QY 1195 TGTCACTCAGCAGAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGGTAAATGAAGAG 1254  
 DB 1397 TGTCACTCAGCAGAAATGCTTTCAGTGTCCAAAGCTTGAAGATC----- 1434  
 QY 1255 AGGATCTCACCCACAGACAAATGCCACATTTTAACTTCTTAAAGAAAGACATCC 1314  
 DB 1435 ----- 1434  
 QY 1315 AGTATCTGAGTCTTTTGAAGACCTGATTAACATCAAGCACTTCCATCATCTTTGT 1374  
 DB 1435 -----ACTGATGAAATTCACACAGTTCATCATCTTTGT 1471  
 QY 1375 CCAAGAAAAATCCTTTTTCATTTGCAAGCCGACACCTCAGATGTAACGTAACAG 1434  
 DB 1472 CTGAGAAAAATCCTTTTTCATTTGCAAGCCGACACCTCAGATGTAACGTAACAG 1531  
 QY 1435 TGGTTTGGGAATGTGAATTAATAGTCTATTAAGAAAACTATGTAATGACAGATC 1494  
 DB 1532 TGGTTTGGGAATGTGAATTAATAGTCTATTAAGAAAACTATGTAATGACAGATC 1591  
 QY 1495 AGCCCAACCGGACTTCCAGGCGCATCCAGATTTGCAAGAGACATCAAAATATGCC 1554  
 DB 1592 AGCCCAACCGGACTTCCAGGACTATCCAGATTTGCC---AGGACAGTTACGAAACGCT 1648  
 QY 1555 TGGACTGAATACAAAGTCAAAAGAAAGATCTGTATGCTTGTATTAATGCACTTGTGAAA 1614  
 DB 1649 TGGACTGAATACAGAGCCGACAAAGATCTGTATCTTGTGCAATGTTCAAGTGAAG 1708  
 QY 1615 CAGCAAAATACATGAATTAATATGATCTGACCTTCAAGTAAACCTGAGTAATCCAA 1674  
 DB 1709 CAGCTGAGTGCATGAATTAATATGATCTGACATCAATAGCTGAGTATCCACAG 1768  
 QY 1675 GAATGTGTTTTGGTCAATTCCTTCTTGTGAACAGACAGACTAGAGGATAGAGCCA 1734  
 DB 1769 GA-----GCCGGGCTTACATTCCTTGTGAACAGAGAAATGAAGTATGAGATCG 1822  
 QY 1735 CCATGGGTTATACGAATCTTACATTAAGAGACTTCACTCGTTGTTGCTCACAGG 1794  
 DB 1823 ACACTGGGTTACAGAAACCTTACATTAAGAGATTAATCTCTCTGATGTCTCACAG 1882  
 QY 1795 TTAACCAATACAGACAGAAACCAAAAGGCTGTGTGTGACATCTTGTATCAGAGAG 1854  
 DB 1883 TTAACCAATACAGACAGAAACCAAAAGGCTGTGTGTGACATCTTGTATCAGAGAG 1942  
 QY 1855 GTGTGTGTGAGCTTGTAAAGAGTATGATCTCAAGAAATATGTGAAGAGTTCTTCAG 1914

Db 1943 GTGTGTGTGAGCTTCTGAGAGAGTGTGGCTGTAAGATATGTGAAGAAGTGTTCAG 2002  
 Qy 1915 ATATCTAGTGAATGAATACATCACTATTTATTCACAAATGAGTGTAGCTTTCTT 1974  
 Db 2003 ATATGAGAGTGAAGGAGTATGATCACTGTATTTACCCCAAGATGGAAGAGCTTTCT 2062  
 Qy 1975 CTGTGCTGATGAGCAACCTCAGCTTCTGACACATCAGAGGATACAGCTTTGACAAATTA 2034  
 Db 2063 CTGTGCTGAGAGCTTCCCTGCTTCTGACACATCAGAGGATACAGCTTTGACAAATTA 2122  
 Qy 2035 CCAGAAATATCTGCGAGAAATATCAATATGCTTCCAGGTTTGTACAGCTTTGATGATCT 2094  
 Db 2123 CCAGAAATATCTGCGAGAAATATCAATATGCTTCCAGATTCATTCAGCTTTGATGATCT 2182  
 Qy 2095 AAATCTCCCAAAATCACTTATTTTACAGATATGCTTAAATGCAATTTTGTATGAGAAATCT 2154  
 Db 2183 AAATCTCCCAAAATCACTTATTTTACAGATATGCTTAAATGCAATTTTGTATGAGAAATCT 2242  
 Qy 2155 CCGTGTGCTGATTTTGAAGTTTGTGTTTATGATGAGGTTAAATATACAAACAGAAAGAT 2214  
 Db 2243 CCGTGTGCTGATTTTGAAGTTTGTGTTTATGATGAGGTTAAATATACAAACAGAAAGAT 2302  
 Qy 2215 TTCACTGAGTGAATGAAAGAGAGGAGATCTTACACTTTAAAGTGAAGTGAAGT 2274  
 Db 2303 TTCACTGAGTGAATGAAAGAGAGGAGATCTTATTTTAAAGTGAAGTGAAGT 2362  
 Qy 2275 AATAGCTGAAAGAGAGATTAATATGATATGAGACATGCTTATGAGGTTGATGCTAT 2334  
 Db 2363 ACCAGCTGAAAGAGAGATTAATATGATATGAGACATGCTTATGAGGTTGATGCTAT 2422  
 Qy 2335 TGTATTAGAGCTGGAATCCATTAATTTTCAAGAGAGAGAGAGAGAGAGTCTCCCTT 2394  
 Db 2423 TGTATTAGAGCTGGAATCTTAAATCTTGAAGAGAGAGAGAGAGAGAGTCTTCAATTC 2482  
 Qy 2395 TTCCCAATATATGAG 2454  
 Db 2483 TTCCCAATATATGAG 2542  
 Qy 2455 CTTCTCTTCTGTGAGTCAAAATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2514  
 Db 2543 CTTCTCTTCTGTGAGTCAAAATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2599  
 Qy 2515 ATGATAGAGTGTCTTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2574  
 Db 2600 GTGAATAGTGTCTTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2659  
 Qy 2575 GAGGAGCTTGTCTTCAACAG 2634  
 Db 2660 GAGGAGCTTGTCTTCAACAG 2708  
 Qy 2635 GATTGTCTTCTTCAACAG 2694  
 Db 2709 ----GTCTCTCTTCAACAG 2764  
 Qy 2695 GCTACAGAGTGAAGTGAAG 2754  
 Db 2765 GCTACAGAGTGAAGTGAAG 2824  
 Qy 2755 GTGACAG 2814  
 Db 2825 GTGACAG 2884  
 Qy 2815 GAG 2874  
 Db 2885 GAG 2944  
 Qy 2875 CTTTGTGAGTGTCTTCAACAG 2913  
 Db 2945 CTTTGTGAGTGTCTTCAACAG 2983

RESULT 9  
 BC026785

**LOCUS** BC026785 3473 bp mRNA linear ROD 06-OCT-2003  
**DEFINITION** Mus musculus serine/threonine kinase 18, mRNA (CDNA clone MGC:30237 IMAGE:5150173), complete cds.  
**ACCESSION** BC026785  
**VERSION** BC026785.1 GI:20072408  
**KEYWORDS** MGC.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**AUTHORS** Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshilyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richardson, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherbina, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.B., Schnerch, A., Schein, J.E., Jones, S.J., and Matre, M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
**TITLE** human and mouse cDNA sequences  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
**MEDLINE** 22388257  
**PUBMED** 12477932  
**REFERENCE** 2 (bases 1 to 3473)  
**AUTHORS** Strausberg, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (04-Apr-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
**COMMENT** Contact: MGC help desk  
 Email: [cgabs-help@mail.nih.gov](mailto:cgabs-help@mail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louieged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.  
**FEATURES**  
 source  
 1..3473  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="MGC:30237 IMAGE:5150173"  
 /tissue\_type="mammary tumor. WAP-10F alpha model. 7 months old, gross tissue."  
 /clone\_11b="NCI CGAP Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLND ac: <http://image.llnl.gov>  
 Series: IRAX Plate: 41 Row: P Column: 11  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27501453.  
 Location/Qualifiers  
 1..3473  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="MGC:30237 IMAGE:5150173"  
 /tissue\_type="mammary tumor. WAP-10F alpha model. 7 months old, gross tissue."  
 /clone\_11b="NCI CGAP Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

```
gene
1..3473.
/ gene="Stk18"
/ note="synonym: 1700028H20"
/ db_xref="locusid:20873"
/ db_xref="mgi:101783"
214..2991
/codon_start=1
/product="serine/threonine kinase 18, isoform sak-a"
/protein_id="AA026785.1"
/db_xref="gi:20072409"
/db_xref="locusid:20873"
/translation="MAACIGERLEDEKRVGNLKGSPAGVYRAESITHTGLEVALIKMID
KKMYRAGVQVQVQNEVKIHQOLKPSVLELYNYPEDNNVYVYLMECHNGENRVLK
NRKPPSEREAHRPHMIOITGMLYLHSHGILHRDLTSLNLTFRNNIKLADGATLO
LNMHPKHVTLCTPNVISPETLRSAHGLEDISGLGCEFTLLGLGRPPIDPTKN
TLNKVLADYEMPAFLSREKODLIHOLLRNPAURLSLSSVLDPHMSNRNPPKSDV
GTVEDMSDGHATLSTTTTASSTGSLGDLBRLLVCPLEPKLITVFQKNSDPS
SGGSAFTCKMCPBOEANSRGRVYLEDERPHRYLRRAHSDRASPNQSRKAT
YSVERCHSVEMLSKPRSLDENQSHNHCLGTPPPADOTPOEMVQWQENLQMN
AHLEGTNEHTVSPNPDYDPLQTLNNAWDTASKNADTSANVNAVYKOLSAWKY
MSAHHKPEVMPQEPDLHPHSEQSKNRSMTGLYQKPLRLSTISPLIARLKPTRK
TKKAVVSLDSEECVELRECESEGVREVIQISDGMTVYVYVYNDGRGPLADRP
PLPTNNISRYSPNLTPEKYNRYQVARSPIQVLRSTPKITTFTRYAKCLMNSPGA
DEWVRYDCAKHKTENLHIIEKTIISTNLKNEKNSIPALSPVDPSCCKGBQASRLS
LSIESVISEEKRSRGSFPLIVGRKPNSTSPRALISLPDPSCKGBQASRLS
VNSAAPTQSPGLSPETVVEG/GHTATATGTGVSLSLPSKADLKSIVKNVWATO
LTSQAVVQFNDGSLVVOAGVSSISYTSPPDQTRRYGENEKLPEYIKQKCLSLIL
LMFSNTPNQ"
247..1008
/ note="S. Tkc, Region: Serine/Threonine protein kinases,
catalytic domain"
/ db_xref="cd:smart00220"

misc_feature
Query Match 64.3%; Score 1874.2; DB 10; Length 3473;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 413; Indels 147; Gaps 9;

ORIGIN
1 ATGGCGACTGTCATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTGTA
214 ATGCGCGCTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTGTA
61 GGATCATTTGCTGCTGCTACAGAGCTGATTCATTCACATGCTTTGGAATTCGCAATC
274 GATCATTTGCTGCTGCTACAGAGCTGATTCATTCACATGCTTTGGAATTCGCAATC
121 AAAATGATGATGAGAAAGCATGTACAAGCAGAGATGTGACAGAGATCCAAATGAG
334 AAAATGATGATGAGAAAGCATGTACAAGCTGGAATGTGACAGAGATCCAAATGAG
181 GTGAAATATCATTCGCAATTTGAAACATCTCTTATCTTGAGCTTTATTAATTTGAA
394 GTGAAATATCATTCGCAATTTGAAACATCTCTTATCTTGAGCTTTATTAATTTGAA
241 GATACCAATTATGTATCTGCTATTTAGAAATGTCGCAATATGAGGAATGAAACAGAT
454 GATACCAATTATGTATCTGCTATTTAGAAATGTCGCAATATGAGGAATGAAACAGAT
301 CTTAAAGATGAGTGAACCTCTTCTCAGAAAAATGAAGCTGCACTTCATGACCAAGATC
514 CTGAAAGACAGATGAGACCTTTCTCAGAAAAAGGAAAGCTAGGCACTTCATGACCAAGAT
361 ATCAGAGGATGTGTATCTTCATTCATGATGATATCAACACCGGACCTTCACTTTCT
574 ATCAGAGGATGTGTATCTTCATTCATGATGATATCAACACCGGACCTTCACTTTCT
421 AACCTCTACTGACTCTGATATATGAACATCAAGATTGCTGATTTTGGCTGGCAACTCA
480 AACATCTTACTTACGCGGAAATATGAACATTAATTTGCTGACTTTGACTAGCAACGAG
634 AACATCTTACTTACGCGGAAATATGAACATTAATTTGCTGACTTTGACTAGCAACGAG
481 CTGAAATATGCAATGAAAGCATATACATTAATGTGAACTCTTAACTATTCACCA
540 TTGATATATGCAATGAAAGCATATACATCTGTGGAAGCTCTTAATATATATTCACCA
694 TTGATATATGCAATGAAAGCATATACATCTGTGGAAGCTCTTAATATATATTCACCA
753
```

```
QY 541 GAAATTCGCACTCGAAGTGCATGAGCCCTTGATCTGATGTTTGGTCCCTGGGCTGTATG 600
DB 754 GAAATTCGCACTCGAAGTGCATGAGCCCTTGATCTGATGTTTGGTCCCTGGGCTGTATG 813
QY 601 TTTTATATCATTTACTATATGCGGAGACCACTTCGACATGACACAGTCAAGAACATTA 660
DB 814 TTTTATATCATTTACTATATGGAAGACCACTTTTGAACATGACACAGTCAAGAACATTA 873
QY 661 AATTAAGATATATGCGAATATGAAATGCAATCTTTTGTCAATAGAGCCCAAGAGC 720
DB 874 AATTAAGATATGCTGCGAATATGAAATGCAATCTTTTGTCAACAGAGCCCAAGAGC 933
QY 721 CTTATATCCAGAGTATCTTCGTAAGAAATCAGACAGATGCTTTAAGTCTGTCTTCAATATG 780
DB 934 CTTATATCCAGAGTATCTTCGTAAGAAACCTTCAGATCGGTTAAGTCTGTCTGTGTG 993
QY 781 GACCATCTTTTATATGCTCCGAAATCTTTCACAAAGATTAAGATTAAGAACTGTGAA 840
DB 994 GACCATCTTTTATATGCTCCGAAATCTTTCACAAAGATTAAGATTAAGAACTGTGAA 1053
QY 841 GACTCAATGATAGTGGGATGCAATTTCTACTGCAATTAAGCTTCTTCCAGTAC 900
DB 1054 GACTCAATGATAGTGGGATGCTACTACTTCCAAACAAATTAAGCTCTTCTGTATAC 1113
QY 901 AGTAAAGTGTATGTTATTTGACAAAGAAAGACTTTGATGTCAGCCACTCCAAAT 960
DB 1114 AGTAAAGTGTACCTTACTTAC--AGAGACTTTGTTGTGCAACCACTTCCAAAT 1170
QY 961 AAAATGACTGATTTCCAAAGATTAAGATTAAGTCAATGATTTTCTTCTGAGAGATGGA 1020
DB 1171 AAAATGACTGATTTCCAAAGATTAAGATTAAGTCAATGATTTTCTTCTGAGAGATGGA 1227
QY 1021 AACGTTTATATCTCAGTGGGAAAT-----CAAGAAACAGTAAATAGTGAGGGGA 1074
DB 1228 AGTAAATTTGATCTCATGAGGAAATCGAAGCAAGAACTAATAGTGGGAGCGGGG 1287
QY 1075 AGAGTAATTCAGATGAGAAAGGCAATCTGATATACCTTCTGATAGCTTATTC 1134
DB 1288 AGAGTAATTCAGATGAGAAAGGCAATCTGATATACCTTCTGATAGCTTATTC 1347
QY 1135 TGTATGATCTGACCTTATATGTCAGTCTCAAGCAAAAAATATTAACATGAGAACGA 1194
DB 1348 TGTATGATCTGACCTTATATGTCAGTCTCAAGCAAAAAATATTAACATGAGAACGT 1404
QY 1195 TGTCACTCAGAGAAATGCTTCAAGTCTCAAAAAAGATCAGAGAGAGTGAATGAAG 1254
DB 1405 TGTCACTCAGAGAAATGCTTCAAGTCTCAAAAAAGATCAGAGAGAGTGAATGAAG 1442
QY 1255 AGTAACTCAGCAAGCAAAATGCAATTTTAACTTTTAAAGAAAGACATCC 1314
DB 1443 ----- 1442
QY 1315 AGTAACTCAGCAATTTTAAAGAACTGATATCAATCAAGCACTTCAATCATCTTTGT 1374
DB 1443 -----ACTGATGAAATCAACAGCTTCAATCATCTTTGT 1479
QY 1375 CAGAGAAATCTCTTTTCAATTTGAGACCGGACCTTCAAGCTGAAACCGTAAACAG 1434
DB 1480 CAGAGAAATCTCTTTTCAATTTGAGACCGGACCTTCAAGCTGAAATGAGTACGAG 1539
QY 1435 TGGTTTGGGAATGCAAAATGCTCAATTAAGAAAACTACTGAATATGACAGATC 1494
DB 1540 TGGTTTGGGAATGCAAAATGCTCAATTAAGAAAACTACTGAATATGACAGATC 1599
QY 1495 AGCCCAAAACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAAATGCC 1554
DB 1600 AGCCCAAAACCGGACCTTCCAGGGCCATCCAGATTTGCAAGATTTGCC---AGGACCTTAAAGAAACGCT 1656
QY 1555 TGGACTGATTAAGAAAGTCAAAAGAACTGATGCTTGTGATATGCAATCTGTAATA 1614
DB 1657 TGGACTGATTAAGAAAGTCAAAAGAACTGATGCTTGTGATATGCAATCTGTAATA 1716
```

1615 CAGCAAAATACATGAAATATATGACTGCACTTCAAGTAACCTGAGTAATCAACA 1674  
 1717 CAGCTGAGTCCATGAAATACATGAGTGCATCACTAAGCTAGGTCATGCCACAG 1776  
 1675 GAATGTTTTTGGCTCAGATCTCTTTTGAACAGAGCAAGATAGGGGTATGAGCCA 1734  
 1777 GA-----GCCGAGACTTACATCTCTCAATTCGAAACAAAGCAAGATAGAGTGGAGTGG 1830  
 1735 CCAATGGGTTTACGAATGCTACATTAAGAGCATTAATCCCGTTGGTGGTCCACAG 1794  
 1831 AACTGGGTTTACGAAGAACTTACCTTAAGAAATATACATCTCTGATGCTCCACA 1890  
 1795 TTAATAACCAATGAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1854  
 1891 TTAAGCCCAATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1950  
 1855 GTGTGTGTGAGCTGTGAAGAGATGCAATCTCAAGAAATATGTGAAGAGTCTTCAG 1914  
 1951 GTGTGTGTGAGCTGTGAAGAGATGCAATCTCAAGAAATATGTGAAGAGTCTTCAG 2010  
 1915 ATATCTAGTATGGAATACGATCACTATTTATTCCAATAGTGTGATAGGTTTCTCT 1974  
 2011 ATATCTAGTATGGAATACGATCACTATTTATTCCTGAAATGATGGAAGGCTTCTCT 2070  
 1975 CTGTGATGATGACCACTCCTCCTACTGACCAATCATGATGATGATGATGATGATGAT 2034  
 2071 CTGTGATGATGACCTCCTCCTCCTACTGACCAATCATGATGATGATGATGATGATGAT 2130  
 2035 CCAGAAATATCTGCGGAAATATCAATATGCTTCAGAGTTGTGATGATGATGATGATGAT 2094  
 2131 CCAGAAATATCTGCGGAAATATCAATATGCTTCAGAGTTGTGATGATGATGATGATGAT 2190  
 2095 AAATCTCCCAAAATCTCTATTTTACAAATATGCTTAAATGATGATGATGATGATGATGAT 2154  
 2191 AAAAATCCCAAAATCTCTATTTTACAAATATGCTTAAATGATGATGATGATGATGATGAT 2250  
 2155 CCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2214  
 2251 CCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2310  
 2215 TTCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2274  
 2311 TTAATTCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2370  
 2275 AATAGCTGAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2334  
 2371 ACCAGCTGAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2430  
 2335 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394  
 2431 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2490  
 2395 TTCCCAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2454  
 2491 TTCCCAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2550  
 2455 CCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514  
 2551 CCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2607  
 2515 ATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
 2608 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2667  
 2575 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634  
 2668 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2716  
 2635 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694  
 2717 ----GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2772  
 2695 GCTACACAGTAACTAGTGAAGTGTGTGGGTTCACTTAAATGATGATGATGATGATGATGAT 2754

2773 GCTACACAGTAACTAGTGAAGTGTGTGGGTTCACTTAAATGATGATGATGATGATGATGAT 2832  
 2755 GTGAGGCGAGAGTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2814  
 2833 GTGAGGCGAGAGTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2892  
 2815 GGAGAAATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2874  
 2893 GGAGAAATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2952  
 2875 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2913  
 2953 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2991

RESULT 10  
 BC057940  
 LOCUS  
 DEFINITION  
 Mus musculus serine/threonine kinase 18, mRNA (cDNA clone MGC:65368  
 IMAGE:5323295), complete cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3377)  
 Strausberg, R.L., Feldgould, B.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shemmer, C.M., Schuler, G.D.,  
 Altschul, S.P., Zeeberg, B., Buetow, K.H., Scherf, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uebachs, T.B., Toshimichi, S.,  
 Carninci, P., Prange, C., Rana, S.S., Loggiano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McSwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, B.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.B.,  
 Schermer, A., Schein, J.E., Jones, S.J., and Matra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 MEDLINE  
 PUBMED  
 12477932  
 2 (bases 1 to 3377)  
 Strausberg, R.  
 Direct Submission  
 Submitted (08-SEP-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgi.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nsl.nih.gov](mailto:nisc_mgc@nsl.nih.gov)  
 Akher, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Grante, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Machuro, Q. J., Masello, C., Maskell, B., Mastrian, S. D., McCloskey, J. C.,  
McDowell, J., Pearson, R., Scantripp, S., Thomas, P. J., Touchman, J. W.,  
Taurgeon, C., Vogt, J. L., Walker, M. A., Weherby, K. D., Wiggins, L.,  
Young, A., Zhang, J. H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Series: IRAC Plate: 123 Row: 1 Column: 15  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA g1: 27501453.

FEATURES  
source

1..3377  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:53368 IMAG:5323295"  
/issue\_type="Mammary tumor, C3(1)-Tag model, Infiltrating  
ductal carcinoma, 5 month old virgin mouse."  
/clone\_1db="NCI CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
1..3377

gene

/gene="Sck18"  
/note="synonym: 1700028H20"  
/db\_xref="LocusID:20873"  
/db\_xref="MGI:101783"  
201..2897

CDS

/codon\_start=1  
/product="Sck18 protein"  
/protein\_id="AAH57940.1"  
/db\_xref="GI:3478391"  
/db\_xref="LocusID:20873"  
/translaction="MAACIGERIEDPKVGNLKGKSPAGVRAESIHITGLAEVAKMID  
KAMYKAGVOROVEMAKIHQIKHPSVLELYNEFDNNVYVLYVEMGNGMNVLYK  
NMKPFSEAEAHFMHIOITGMUYLHSHGLHRDLISMLITRMNLIKIDPGALATQ  
LNNPKHYTLGGTTPYISPEIATRSAGHLESDISLGMFTLLIGRPFDITVKN  
TLNKVYLADYEMAPLSRAQDLIHQLRPNADRLJSVLDHPFMSRNPSPKSDV  
GTVESDMSDGHATLSTTTTASGTSLSGLDLDRLLVQPLPKITVQKNNSSDFS  
SEDSNFTCKWNPQOANSRGRVLEDAERPHSRYLRAHSSDRASPSNQSATK  
YSEVERCHSEVMSKPRSLDENQSHNHCLGTPPEPADOTPOKEMVQOMGNLOM  
AHIGTNEHTVSPNRFQDYPDLDTLANAWTDRAGKNADTSANVAIVOLSAMKY  
MSAHHKPEVMEPOEDLHPHSEKSNRSESLGYOKTSLISITSLAHAKLPTRQK  
RLOLVRSKTPKITITPRAKTILMENSADFEVFNWISRSFDMLEKRWKQYAS  
SNTLNKNEVITLKEEVKVMHANEGRICTLSLSVLSBEKRSSTSPVPIIVGRTA  
PENTSSPKALSAIPVDPSCCKGBOASASRLSYNSAAPTOSBEKRSSTSPVPIIVGRTA  
TATGTGSSSLPKSAQLKSYFVKVGNATOLSGAVVWVFQDSQLVVOAGVSISIV  
TSPDQTTTRYGENEKLPEYIKQKLCSSILLMPSNPFPNQ"  
234..995  
/note="S\_TKC; Region: Serine/Threonine protein kinases,  
catalytic domain"  
/db\_xref="CD:smart00220"

ORIGIN

Query Match 59.2%; Score 1724.6; DB 10; Length 3377;  
Best Local Similarity 78.5%; Pred. No. 0;  
Matches 2292; Conservative 0; Mismatches 399; Indels 228; Gaps 10;

QY 1 ATGGGACCTGCTGATGGGAGAGAAATGAGATTTTAAAGTTGGAATCTGCTGTGTA 60  
DB 201 ATGGGGGGGTGATGGGAGAGATCGAGACTTTAAGTTGGAATCTACTCGGTA 260  
QY 61 GGATCATTTGCTGGGTCTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 120  
DB 261 GGATCATTTGCTGGGTCTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 320  
QY 121 AAAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTAAGAGAGTCCAAATAG 180  
DB 321 AAAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTAAGAGAGTCCAAATAG 380  
QY 181 GTGAATATACATTCGCAATGAACATCTTCTATCTTGAAGCTTTATAACTATTTGAA 240

DB 381 GTGAATATACATTCGCAATGAACATCTTCTATCTTGAAGCTTTATAACTATTTGAA 440  
QY 241 GATAGCAATTATGTGATCTGGATATTAAGAAATGTCATTAAGGAATGAACGAT 300  
DB 441 GATAGCAATTATGTGATCTGGATATTAAGAAATGTCATTAAGGAATGAACGAT 500  
QY 301 CTAAAGATATGATGAAGAAAGCCCTTCTCAGAAATGAAGCTCGACACTTCATGACAGATC 360  
DB 501 CTGAAGAACAGATGAAGCTTCTCAGAAAGGAAGCTTGAAGCTTATGACACGAT 560  
QY 361 ATCAGAGGATGTGATCTTCAATCTCATGTATATCTACACCGGACCTCACTTCT 420  
DB 561 ATCAGAGGATGTGATCTTCAATCTCATGTATATCTATGACACCGGACCTCACTTCT 620  
QY 421 AACTCTTACTACTCGTATATATGAATCAAGATTCGATTTTGGCTGGCACTCA 480  
DB 621 AACTCTTACTACTCGGATATATGAATCAAGATTCGATTTTGGCTGGCACTCA 680  
QY 481 CTGAAGAAATGCCATGAAGAAAGCATATACATTAATGGAATCTCTAATCAATTCACA 540  
DB 681 TTGAATATGCAATGAAGAAAGCATATACATCTGTGGACTCTTAATATATTCACA 740  
QY 541 GAAATTCGCACTCGAAGTGCATGAGCTTGAATCTGATGTTGGTCCGCTGATG 600  
DB 741 GAAATTCGCACTCGAAGTGCATGAGCTTGAATCTGATGTTGGTCCGCTGATG 800  
QY 601 TTTTATACATTAATTCGAGAGAACACCTTGCACATGACAGTCAAGAACATTA 660  
DB 801 TTTTATACATTAATTCGAGAGAACACCTTGCACATGACAGTCAAGAACATTA 860  
QY 661 AATAAGATGATTTGGAGATTAAGAAATGCCATCTTTTGTCAATAGAGCCAGAG 720  
DB 861 AATAAGATGATTTGGAGATTAAGAAATGCCATCTTTTGTCAATAGAGCCAGAG 920  
QY 721 CTATATCACAGATTAATTCGAGAGAACACCTTGTGACATGAGTGTGCTGCTG 780  
DB 921 CTATATCACAGATTAATTCGAGAGAACACCTTGTGACATGAGTGTGCTGCTG 980  
QY 781 GACATCTTTTATGTCGCAAAATCTTCAACAAAGATTAAGATTTAGAACTGTG 840  
DB 981 GACATCTTTTATGTCGCAAAATCTTCAACAAAGATTAAGATTTAGAACTGTG 1040  
QY 841 GACTCAATGATATGAGGAGATGCAATTTCTACTGCAATTAACAGCTTCCAGTACC 900  
DB 1041 GACTCAATGATATGAGGAGATGCAATTTCTACTGCAATTAACAGCTTCCAGTACC 1100  
QY 901 AGTATAGTGTGATTTATTTGACAAAGAGACCTTTGATGTCAGCCTCCCAAT 960  
DB 1101 AGTATAGTGTGATTTATTTGACAAAGAGACCTTTGATGTCAGCCTCCCAAT 1157  
QY 961 AAAATGACTGATTTTCAAGAAATGAAGTCAAGATTTTCTCTCAGAGATGGA 1020  
DB 1158 AAAATGACTGATTTTCAAGAAATGAAGTCAAGATTTTCTCTCAGAGATGGA 1214  
QY 1021 AACAATTTTATATCTAGTGGGAAAT-----CAAGAAACAGATTAATGTAAGAGGGA 1074  
DB 1215 AATTAATTTTATATCTAGTGGGAAATCAAGAAACAGATTAATGTAAGAGGGA 1274  
QY 1075 AGAGTAATTCAGATTCAGAAAGAGCCACATTTCTGATTAACCTTGTAAGACTTATTC 1134  
DB 1275 AGAGTAATTCAGATTCAGAAAGAGCCACATTTCTGATTAACCTTGTAAGACTTATTC 1334  
QY 1135 TCTGATAGTCTGGCACTTCTATATGTCAGTCTCAAGCAAAACATATCAATGAGGA 1194  
DB 1335 TCTGATAGTCTGGCACTTCTATATGTCAGTCTCAAGCAAAACATATCAATGAGGA 1391  
QY 1195 TGTCACTCAGCAAAATGCTTTCAAGTCTCAAGATTCAGAGAGGAGGTAATGAG 1254  
DB 1392 TGTCACTCAGTGAATGCTTTCAAGATTCAGAGATTC----- 1429  
QY 1255 AGGTACTCACCACAGACAAATGCAATTTTAACTTCTTAAAGAAAGACATCC 1314

1430 ----- 1429  
 1315 AGTACTTGGATCTTTGAAAAGCCTGTATTAACATCAAGCACTCTCCATCATCTTTGT 1374  
 1430 -----ACTGATGAAAATCAACACAGTTCCATCATCTATTTGT 1466  
 1375 CCAGAAAAACCTCTTTTCATTTTGACAGCCGACACCTCAGACTGAACCGTACACAG 1434  
 1467 CTAGAAAAACCTCTTTTCATTTTGACAGCCGACACCTCAGACTGAACCGTACACAG 1526  
 1435 TGGTTTGGGAATCTGCAATAAATGCTCATTTTAAAGAAAACTACTGAATATGACAGCATC 1494  
 1527 TGGTTTGGGAATCTGCAATAAATGCTCATTTTAAAGAAAACTACTGAATATGACAGCATC 1586  
 1495 AGCCCAACCGGACCTTCCAGGCGCATCCAGATTGCGAAGAGACATCAAAAAATGCGC 1554  
 1587 AGCCCAACCGGACCTTCCAGGCGCATCCAGATTGCGAAGAGACATCAAAAAATGCGC 1643  
 1555 TGGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTGTATTAATGCAATTCGTAA 1614  
 1644 TGGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTGTATTAATGCAATTCGTAA 1703  
 1615 CAGCAAAATACCATGAAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1674  
 1704 CAGTGAAGTGCATGAAATATATGAGTGCATCATCATTAAGCTGAGTCAATGCCACAG 1763  
 1675 GAATGTTTGGTGGCTCAGATCTCTTTCTGAAACAGACAGACTAGGGGTATGAGCCA 1734  
 1764 GA-----GCCGACCTTACATCTCTCAATTCGAAACAAAGCAAGATGAAAGTGGAGTGG 1817  
 1735 CATGGGGGTTATCAGAAATGATATTAAGAGCATTAATCCGTTGGTGGTCCACAG 1794  
 1818 ACCTGGGTTATCAGAAATGATATTAAGAGCATTAATCCGTTGGTGGTCCACAG 1877  
 1795 TTAATAACCAATGACAGCAAAAAACCAAAAAGCTGTGGTGAAGTACTTGAATTCAGAGAG 1854  
 1878 TTAAGGCAATGACAGCAAAAAACCAAAAAGCTGTGGTGAAGTACTTGAATTCAGAGAG 1936  
 1855 GTGTGTGGAGCTGTGTAAGAGATGCACTCAAGAAATATGTAAAGAGTTCTTCAG 1914  
 1937 ----- 1936  
 1915 ATATCTAGATGAGAAATACGATCACTATTATTAATCAAAATGATGATGAGTTTCC 1974  
 1937 -----GATCAGCTGTTTATTAATCCGAAACGATGAGAGGCTTTCC 1976  
 1975 CTGTGATAGACCAACCTCAGCTACTGACCAACATCAGTAGTACAGCTTTGCAATTTA 2034  
 1977 CTGTGATAGACCAACCTCAGCTACTGACCAACATCAGTAGTACAGCTTTGCAATTTA 2036  
 2035 CCAGAAAAATCTGCGGAAATATGCAATATGCTTCCAGTTGTACAGCTTGAATCT 2094  
 2037 CCAGAAAAATCTGCGGAAATATGCAATATGCTTCCAGTTGTACAGCTTGAATCT 2096  
 2095 AATCTCCCAAAATCACTATTATTAAGATATGCAATATGCTTCCAGTTGTACAGCTTGAATCT 2154  
 2097 AATCTCCCAAAATCACTATTATTAAGATATGCAATATGCTTCCAGTTGTACAGCTTGAATCT 2156  
 2155 CCGTGTGCTGATTTTGAAGTTGGTTTATGATGGGTAATAAATCACAAAACAGAAAT 2214  
 2157 CCGTGTGCTGATTTTGAAGTTGGTTTATGATGGGTAATAAATCACAAAACAGAAAT 2216  
 2215 TTCAATTCAGGTGATTAAGAAAGACAGGAAGTCTTACCTTTAAAGTGAAGTT 2274  
 2217 TTCAATTCAGGTGATTAAGAAAGACAGGAAGTCTTACCTTTAAAGTGAAGTT 2276  
 2275 AATAGCTTGAAGAGAGATTAATAATGTATATGAGCAATGCTATAGGGTCACTGAT 2334  
 2277 ACAGGCTTGAAGAGAGATTAATAATGTATATGAGCAATGCTATAGGGTCACTGAT 2336  
 2335 TGTATGACATGCAATTCATTAATTTGACAGAGAAAAAGAAATCTAGAGTCTCCCTTT 2394  
 2337 TGTATGACATGCAATTCATTAATTTGACAGAGAAAAAGAAATCTAGAGTCTCCCTTT 2396

2395 TTCCCAATATCATAGGAAGAAAACTGGTAGTACTAGTTCACTTAAGGCTTATCACCT 2454  
 2397 TTCCCAATATCATAGGAAGAAAACTGGTAGTACTAGTTCACTTAAGGCTTATCACCT 2456  
 2455 CCTCTCTTGTGGATTCAAATTAACCAACGAGAGATAGAGCATCTTTCAACGAATGGTC 2514  
 2457 CCTC---CTGTGAGCCCAAGCTGTATAGGAGAGAGAGGCGCTCAGCAAGACAGCTGAGC 2513  
 2515 ATGCATAGTGTCTGCTTTCTTCAACAGAGCAACCAATCCTTAATCCCTATGTTCAAT 2574  
 2514 GTGAATAGTGTCTGCTTTCTTCAACAGAGCAACCAATCCTTAATCCCTATGTTCAAT 2573  
 2575 GAAGGACTGTGTCTTACAACTACAGCTTCTGGAACAGACATCTCTTCAATAGTCAAAA 2634  
 2574 GAAGGACTGTGTCTTACAACTACAGCTTCTGGAACAGACATCTCTTCAATAGTCAAAA 2622  
 2635 GATGTCTTCTTCAATTCAGACCAACTTTTGAATCTGTTTGTGAATAATGTTGGTGG 2694  
 2623 ----GCTTCTTAATCTGCAACAGCTTTGAATCTGTTTGTGAATAATGTTGGTGG 2678  
 2695 GCTACACAGTTAACTATAGTGAAGCTGTGTGGTTCAGTTTAAAGATGGTCCAGTTGGT 2754  
 2679 GCTACACAGTTAACTATAGTGAAGCTGTGTGGTTCAGTTTAAAGATGGTCCAGTTGGT 2738  
 2755 GTGACAGCAGAGTGTCTTCAATTCAGTTTAACTTCACTCAACCAATGTCACCAACTAGAT 2814  
 2739 GTGACAGCAGAGTGTCTTCAATTCAGTTTAACTTCACTCAACCAATGTCACCAACTAGAT 2798  
 2815 GGAGAAATGAATAATTAACCAAGTATCAATCAACGAATAATTAAGTGTCTTCCATC 2874  
 2799 GGAGAAATGAATAATTAACCAAGTATCAATCAACGAATAATTAAGTGTCTTCCATC 2858  
 2875 CTTTGAATGTTTCTTAATCCGACTGCTTAATTTTCATGA 2913  
 2859 CTTTGAATGTTTCTTAATCCGACTGCTTAATTTTCATGA 2897

RESULT 11  
 BC060363  
 LOCUS 3554 bp mRNA linear VMT 21-OCT-2003  
 DEFINITION Xenopus laevis cDNA clone MGC:68791 IMAGE:4202704, complete cds.  
 ACCESSION BC060363  
 VERSION BC060363.1 GI:37805196  
 KEYWORDS MGC.  
 SOURCE  
 ORGANISM  
 Xenopus laevis (African clawed frog)  
 Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 3554)  
 Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.  
 and Richardson, P.  
 Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 Initiative  
 JOURNAL  
 MEDLINE  
 PUBMED  
 22341132  
 12454917  
 2 (bases 1 to 3554)  
 Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Ramet, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshitake, S.,  
 Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Millar, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,





REFERENCE 1 (bases 1 to 57057)  
AUTHORS Sulston,J.B. and Waterston,R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074

REFERENCE 2 (bases 1 to 57057)  
AUTHORS Isak,A., Kozlowski,A. and Creason,K.  
TITLE The sequence of Homo sapiens BAC clone RP11-398H1  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (14-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (21-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 57057)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 16, 2002 this sequence version replaced gi:1885145.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: sapiens@watson.wustl.edu  
----- Summary Statistics  
Center project name: H\_NH0398H01  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Moon,P.Y., Zhao,B., Frengen,B., Tatenno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries: Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org  
VECTOR: pBace3.6

NEIGHBORING SEQUENCE INFORMATION:

FEATURES A single plasmid subclone region exists between bases 27391-27445.  
Source The clone sequenced to the left is RP11-50D9, 2000 bp overlap; the clone sequenced to the right is RP11-214014, 2000 bp overlap. Actual start of this clone is at base position 115600 of RP11-50D9.

1..57057  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-398H1"  
/clone\_11b="RPc1-11"  
1..127  
repeat\_region  
/rpt\_family="ERV1"  
129..246  
repeat\_region  
/rpt\_family="L1"  
247..532  
repeat\_region  
/rpt\_family="Alu"  
533..587  
repeat\_region  
/rpt\_family="AT-rich"  
761..1075  
repeat\_region  
/rpt\_family="L1"  
1106..1410  
repeat\_region  
/rpt\_family="L2"  
1426..1717  
repeat\_region  
/rpt\_family="Alu"  
1817..2029  
repeat\_region  
/rpt\_family="L1"  
2169..2217  
repeat\_region  
/rpt\_family="U5"  
2316..2359  
repeat\_region  
/rpt\_family="CATRA)n"  
2534..2622  
repeat\_region  
/rpt\_family="ERV1"  
2623..2926  
repeat\_region  
/rpt\_family="Alu"  
2927..3615  
repeat\_region  
/rpt\_family="ERV1"  
3723..4337  
repeat\_region  
/rpt\_family="ERV1"  
4510..4672  
repeat\_region  
/rpt\_family="MaLR"  
4736..7003  
repeat\_region  
/rpt\_family="L1"  
7004..7098  
repeat\_region  
/rpt\_family="L1"  
12632..12658  
repeat\_region  
/rpt\_family="AT-rich"  
12659..12998  
repeat\_region  
/rpt\_family="L1"  
13006..13923  
repeat\_region  
/rpt\_family="L1"  
13924..14216  
repeat\_region  
/rpt\_family="Alu"  
14217..15190  
repeat\_region  
/rpt\_family="L1"  
15251..15287  
repeat\_region  
/rpt\_family="(TG)n"  
15363..15631  
repeat\_region  
/rpt\_family="L1"  
15644..15938  
repeat\_region  
/rpt\_family="ERV1"  
15939..15966  
repeat\_region  
/rpt\_family="AT-rich"  
16041..16344  
repeat\_region  
/rpt\_family="ERV1"  
16480..16874  
repeat\_region  
/rpt\_family="L1"  
16884..17266  
repeat\_region  
/rpt\_family="MER1\_type"  
17518..18199  
repeat\_region

```

repeat_region      18294..18407      /rpe_family="L2"
                    /rpe_family="MIR"
repeat_region      18703..19006      /rpe_family="Alu"
repeat_region      19096..19359      /rpe_family="Alu"
repeat_region      19376..19395      /rpe_family="Alu"
repeat_region      19415..19715      /rpe_family="(TAA)n"
repeat_region      20200..20308      /rpe_family="Alu"
repeat_region      20309..21114      /rpe_family="MALR"
repeat_region      21115..21214      /rpe_family="L1"
repeat_region      21351..21658      /rpe_family="Alu"
repeat_region      21676..22978      /rpe_family="Alu"
repeat_region      23115..23137      /rpe_family="L1"
repeat_region      23146..23168      /rpe_family="AT_rich"
repeat_region      23421..23513      /rpe_family="AT_rich"
repeat_region      23533..23827      /rpe_family="L1"
repeat_region      23833..24130      /rpe_family="Alu"
repeat_region      24195..24494      /rpe_family="Alu"
repeat_region      24517..24684      /rpe_family="Alu"
repeat_region      24698..24747      /rpe_family="L1"
repeat_region      24787..25165      /rpe_family="AT_rich"
repeat_region      25229..25315      /rpe_family="L1"
repeat_region      25328..25387      /rpe_family="MER2_type"
repeat_region      25426..25733      /rpe_family="MER2_type"
repeat_region      /rpe_family="Alu"

Query Match      35.2%; Score 1025; DB 9; Length 57057;
Best Local Similarity 98.1%; Pred. No. 8.8e-213;
Matches 1037; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

```

QY      696 TTTTGTGATAGAGGCGCAAGGACCTTATTCACGAGTACTTCGTGAGAAATCCAGACA 755
DB      33010 TTTTGTGATAGAGGCGCAAGGACCTTATTCACGAGTACTTCGTGAGAAATCCAGACA 33069
QY      756 TCGTTAAGTCTGTCTTCAGATATGGACCATCTTTATATGTCGGAAATTTTCAACAA 815
DB      33070 TCGTTAAGTCTGTCTTCAGATATGGACCATCTTTATATGTCGGAAATTTTCAACAA 33129
QY      816 AAGTAAAGTTTAAAGAACTGTGGAAGACTCAATATGATATGAGGCAAGCCAAATTTCTAC 875
DB      33130 AAGTAAAGTTTAAAGAACTGTGGAAGACTCAATATGATATGAGGCAAGCCAAATTTCTAC 33189
QY      876 TGCATTAAGCTCTTCAGATACAGATTAAGTGTATTTATTTGACAAAAGAGACT 935
DB      33190 TGCATTAAGCTCTTCAGATACAGATTAAGTGTATTTATTTGACAAAAGAGACT 33249
QY      936 TTTGATTTGTGACGCCATCCCAATTAATGAATGATTTTCCAAAGATTAAGTTCAAC 995
DB      33250 TTTGATTTGTGACGCCATCCCAATTAATGAATGATTTTCCAAAGATTAAGTTCAAC 33309
QY      996 TGAATTTTCTTCTTCAGAGATGGAACGTTTATATCTAGTGTGGGAAATCAAGAAC 1055
DB      33310 TGAATTTTCTTCTTCAGAGATGGAACGTTTATATCTAGTGTGGGAAATCAAGAAC 33369
QY      1056 CAGTATAGTGGAAAGGGAAGATATTCAGATGCAAGAAAGGCGCAATTCTGATA 1115
DB      33370 CAGTATAGTGGAAAGGGAAGATATTCAGATGCAAGAAAGGCGCAATTCTGATA 33429
QY      1116 CCTTGTAGAGCTTATCTCTGATAGATCTGGCACTTTATATGTCAGTCTTCAGCAA 1175
DB      33430 CCTTGTAGAGCTTATCTCTGATAGATCTGGCACTTTATATGTCAGTCTTCAGCAA 33489
QY      1176 AACATATACATGGAACGATGTCATCAGACAGAAATGCTTTCAGTCTCCAAAAGATCAGG 1235
DB      33490 AACATATACATGGAACGATGTCATCAGACAGAAATGCTTTCAGTCTCCAAAAGATCAGG 33549
QY      1236 AGAGGTGAAAATGGAAGAGGTAATCAACCAAGCAACATGCAATTTTAACTT 1295
DB      33550 AGAGGTGAAAATGGAAGAGGTAATCAACCAAGCAACATGCAATTTTAACTT 33609
QY      1296 CTTTAAAGAAAAGACATCCAGATGTTCTGATCTTTTGAAGACCTGATTAACATCAAGC 1355
DB      33610 CTTTAAAGAAAAGACATCCAGATGTTCTGATCTTTTGAAGACCTGATTAACATCAAGC 33669
QY      1356 ACTCTCAATCAATCTTGTCCAGAAAACTCCTTTT 1392
DB      33670 ACTCTCAATCAATCTTGTCCAGAAAACTCCTTTT 33706

RESULT 13
AL591646      153342 bp      DNA      linear      HTG 15-AUG-2001
LOCUS      Homo sapiens chromosome 1 clone RP11-25411, 3 unordered pieces.
DEFINITION      AL591646
ACCESSION      AL591646
VERSION      AL591646.6 GI:14586307
KEYWORDS      HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1      Melay,K.
   Direct Submision
   Submitted (14-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
   CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
   Request: clonerequest@sanger.ac.uk
   On Jul 2, 2001 this sequence version replaced gi:14330185.
   ----- Genome Center
   Center: Sanger Centre
   Center code: SC
   Web site: http://www.sanger.ac.uk
   Contact: humquery@sanger.ac.uk

```

----- Project Information  
Center project name: ba25411  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; 108752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 153071 bases at least Q40  
Consensus quality: 153095 bases at least Q30  
Insert size: 153142; sum-of-contrigs  
Insert size: 154993; 6.8% error; agarose-fp  
Quality coverage: 11.28x in Q20 bases; sum-of-contrigs  
Coverage: 13.04x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 91998: contig of 91998 bp in length  
\* 91999 92098: gap of 100 bp  
\* 92099 140835: contig of 48737 bp in length  
\* 140836 140935: gap of 100 bp  
\* 140936 153342: contig of 12407 bp in length.  
Location/Qualifiers  
1. .153342  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-25411"  
/clone\_id="RP11-11.1"  
1. .91998  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:SP6  
vector\_side:right"

## FEATURES

source

misc\_feature  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:SP6  
vector\_side:right"

misc\_feature  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:SP6  
vector\_side:right"

misc\_feature  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:SP6  
vector\_side:right"

## ORIGIN

Query Match 35.2%; Score 1025; DB 2; Length 153342;  
Best Local Similarity 98.1%; Pred. No. 7.6e-213; Indels 0; Gaps 0;  
Matches 1037; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
-----  
QY 336 AGCTCGACACTTCATGACCAAGATCATACAGGAGTTGTATCTTCATTCATGGTAT 395  
DB 112321 AGCTCGACACTTCATGACCAAGATCATACAGGAGTTGTATCTTCATTCATGGTAT 112380  
QY 396 ACTACACCGGAGCTTCACACTTTCTAAGCTCTCTACTGACCTGTAATGAAATCAAGAT 455  
DB 112381 ACTACACCGGAGCTTCACACTTTCTAAGCTCTCTACTGACCTGTAATGAAATCAAGAT 112440  
QY 456 TGCATATTTGGGGCTGGCACTCAACTGAAATCCATGAAAGCACTATACATATG 515  
DB 112441 TGCATATTTGGGGCTGGCACTCAACTGAAATCCATGAAAGCACTATACATATG 112500  
QY 516 TGGAACTCTCACTACTATTCACCAAGAAATTCGACTGGAAGTCAATGACCTGTAATC 575  
DB 112501 TGGAACTCTCACTACTATTCACCAAGAAATTCGACTGGAAGTCAATGACCTGTAATC 112560  
QY 576 TGAATGTTGGCTCCTGGGCTGTATGTTTATATCACTTATTCGGGAGACCAACCTTCGA 635  
DB 112561 TGAATGTTGGCTCCTGGGCTGTATGTTTATATCACTTATTCGGGAGACCAACCTTCGA 112620

QY 636 CACTGACACAGTCAAGAACATTAATAAGTAGTATGGCAGATTATGAAATGCCATC 695  
DB 112621 CACTGACACAGTCAAGAACATTAATAAGTAGTATGGCAGATTATGAAATGCCATC 112680  
QY 696 TTTTGTGCAATGAGGCGCAAGACCTTATTCACCAAGTACTTGTGTAAATTCACGAGA 755  
DB 112681 TTTTGTGCAATGAGGCGCAAGACCTTATTCACCAAGTACTTGTGTAAATTCACGAGA 112740  
QY 756 TCGTTAAGTCTGTCTTCAGATATGAGCAATCCTTTATGTCGGAAATCTTCACAAA 815  
DB 112741 TCGTTAAGTCTGTCTTCAGATATGAGCAATCCTTTATGTCGGAAATCTTCACAAA 112800  
QY 816 AAGTAAAGTTAGGAAGTGTGAAAGCTCAATGATAGTGGGACATGACATTTCTAC 875  
DB 112801 AAGTAAAGTTAGGAAGTGTGAAAGCTCAATGATAGTGGGACATGACATTTCTAC 112860  
QY 876 TGCATTAACAGCTTCTTCACGATACCAAGTATAGTGTATTTATTTGCAAAAGAGACT 935  
DB 112861 TGCATTAACAGCTTCTTCACGATACCAAGTATAGTGTATTTATTTGCAAAAGAGACT 112920  
QY 936 TTTGATGTCAGCCACTCCCAATTAATGACTGTATTTGCAAAAGATTAAGTTCAAC 995  
DB 112921 TTTGATGTCAGCCACTCCCAATTAATGACTGTATTTGCAAAAGATTAAGTTCAAC 112980  
QY 996 TGAATTTTCTTCTTCAGAGATGGAACAGTTTATATCAAGTGGGAAATCAAGAAAC 1055  
DB 112981 TGAATTTTCTTCTTCAGAGATGGAACAGTTTATATCAAGTGGGAAATCAAGAAAC 113040  
QY 1056 CAGTAAATGTGGAAGGGGAGAGATTAATTCAGATGCAAGAAAGGCAATTCGATA 1115  
DB 113041 CAGTAAATGTGGAAGGGGAGAGATTAATTCAGATGCAAGAAAGGCAATTCGATA 113100  
QY 1116 CTTGTGTAGCTTATTCCTCTGATAGATCTGCACTTCTAATAGTCACTCAAGCAA 1175  
DB 113101 CTTGTGTAGCTTATTCCTCTGATAGATCTGCACTTCTAATAGTCACTCAAGCAA 113160  
QY 1176 AACATATACATGGAAGAGATGCTACTCAGCAGAAATGCTTTCAGTCCAAAGATCAGG 1235  
DB 113161 AACATATACATGGAAGAGATGCTACTCAGCAGAAATGCTTTCAGTCCAAAGATCAGG 113220  
QY 1236 AGGAGGTGAAATGGAAGAGGATCACTCAACCCACAGAACAAATGCGCAATTTTAACTT 1295  
DB 113221 AGGAGGTGAAATGGAAGAGGATCACTCAACCCACAGAACAAATGCGCAATTTTAACTT 113280  
QY 1296 CTTTAAAGAAAGACATCAATGATCTGATCTTTGAAAGACCTGATTAACATCAAGC 1355  
DB 113281 CTTTAAAGAAAGACATCAATGATCTTGTGATCTTTGAAAGACCTGATTAACATCAAGC 113340  
QY 1356 ACTCTCAATCATCTTGTCCAGAAAACCTCTTTT 1392  
DB 113341 ACTCTCAATCATCTTGTCCAGAAAACCTCTTTT 113377  
-----  
RESULT 14  
AR083570 1600 bp DNA linear PAT 01-SEP-2000  
LOCUS AR083570 Sequence 5 from patent US 5976893.  
DEFINITION AR083570  
ACCESSION AR083570  
VERSION AR083570.1 GI:10010343  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1600)  
AUTHORS Dennis J.W., Hefferman M. and Fode C.  
TITLE Methods for identifying binding partners, agonists, and antagonist  
of a serine/threonine tyrosine kinase  
JOURNAL Patent: US 5976893-A 5 02-NOV-1993;  
FEATURES Location/Qualifiers  
1. .1600  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 33.3%; Score 970.4; DB 6; Length 1600;  
 Best Local Similarity 85.0%; Pred. No. 1.1e-200;  
 Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

1 ATGGGCACTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTGTA 60  
 206 ATGGGGCGCTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTGTA 265  
 61 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 120  
 266 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 325  
 121 AAAATGATGATPAAGAAAGCATGTACAAAGCAGAAAGTGTACAGAGTCCAAATGAG 180  
 326 AAAATGATGATPAAGAAAGCATGTACAAAGTGTACAGAGTCCAAATGAG 385  
 181 GTGAAAATACATTTGCCAATTTGAAACATCCCTTATCTTGAGCTTTATTAATTTTGA 240  
 386 GTGAAAATACATTTGCCAATTTGAAACATCCCTTATCTTGAGCTTTATTAATTTTGA 445  
 241 GATAGCAATTAATGTATCTGTGATTAAGAAATGTCCATTAATGAGAAATGAACAGTAT 300  
 446 GATTAACAATTAATGTATCTGTGATTAAGAAATGTCCATTAATGAGAAATGAACAGTAT 505  
 301 CTAAAGAAATGAGTGAACCCCTTCTGAGAAATGAAGCTCGACATTCATGACCGAGATC 360  
 506 CTGAAGAAACAGATGAAGACCTTTCTGAGAAAGGAAAGCTTAGAGCTTCATGACCGAGAT 565  
 361 ATCAGAGGATGTGTATCTTCTCATTTCTCATGTATCTACACCGGAGCTTCACATTTTC 420  
 566 ATCAGAGGATGTATCTTCTCATTTCTCATGTATCTACACCGGAGCTTCACATTTTC 625  
 421 AACCTCTACGACTCGTATTAATGAACATCAAGATGCTGATTTGGCTGGCACTCA 480  
 626 AACATCTTACTTACCGCGAATATGAACATTAATTTGCTGACTTGAACAGGACG 685  
 481 CTGAAGAAAGCAGTGAAGGACATTAATTAATGGAACCTCTTAATCAATTCACCA 540  
 686 TTGAATATGCAATGAAGGACATTAATTAATGGAACCTCTTAATTAATTCACCA 745  
 541 GAAATGGCACTGAGAGTGCACATGCGCTTGAATCTGATGTTGGTCCCTGGGCTGTATG 600  
 746 GAAATGGCACTGAGAGTGCACATGCGCTTGAATCTGATGTTGGTCCCTGGGCTGTATG 805  
 601 TTTTATACATTAATTAATGAGAGCACCCTTTCGACATGACACAGTCAAGAACATTA 660  
 806 TCTTATACATTAATTAATGAGAGCACCCTTTCGACATGACACAGTCAAGAACATTA 865  
 661 AATAAAGTATGATGAGCAATTAAGAAATGCGCATTTTGTTCATAGAGCCAGAGAC 720  
 866 AATAAAGTATGATGAGCAATTAAGAAATGCGCATTTTGTTCATAGAGCCAGAGAC 925  
 721 CTTATTCACAGTTACTTGTGTAAGAAATTCAGCAGATCGTTTAAAGTCTGCTCAGATTTG 780  
 926 CTTATTCACAGTTACTTGTGTAAGAAATTCAGCAGATCGTTTAAAGTCTGCTCAGATTTG 985  
 781 GACCATCTTTTATGTCGGAATCTTTCACAAAAGATTAAGATTAGGAACCTGTGAA 840  
 986 GACCATCTTTTATGTCGGAATCTTTCACAAAAGATTAAGATTAGGAACCTGTGAA 1045  
 841 GACTCAATTTGATAGTGGGCAATGCAATTTTCTACGCAATTAAGCTTTCCAGTATCC 900  
 1046 GACTCAATTTGATAGTGGGCAATGCAATTTTCTACGCAATTAAGCTTTCCAGTATCC 1105  
 901 AGTATTAAGTATGATTTATTTGCAAAAGAAAGCTTTGATTTGTCAGCACTCCCAAT 960  
 1106 AGTATTAAGTATGATTTATTTGCAAAAGAAAGCTTTGATTTGTCAGCACTCCCAAT 1162  
 961 AAAATGATGATTTTCAAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020  
 1163 AAAATGATGATTTTCAAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1219

1021 AACAGTTTATTAATCACTGAGGGGAAT-----CAAGAAACAGTAATAGTGAAGGGA 1074  
 1220 AGTAATTTTGTACTCAATGGGGAATTCAGAAACAAAGAGCTTAATGAGGGGACGGGG 1279  
 1075 AGAGTAATTCAGAGTCAAGAAAGGAGGAGCAATTCGTATGATCTTGTAGAGTTATCC 1134  
 1280 AGAGTAATTCAGAGTCAAGAAAGGAGGAGGAGCAATTCGTATGATCTTGTAGAGTTATCC 1339  
 1135 TCTGATAGATCTGAGCACTTCTTAATAGTCTCAAGCAAAAAATATCAATGAAAGCA 1194  
 1340 TCTGATAGAGCCAGCCCTCTTA--TCAGTCTCGAGCAAAAAATATCAATGAAAGCA 1396  
 1195 TGTCACTGAGCAAAATGCTTTCAGTGTCCAAAGATCGAGAGGTGAAAAATGAAGAG 1254  
 1397 TGTCACTGAGTAATGCTTTCAGAGCTTCAAGCTTAAGATC--ACTGATGAAATCAAC 1453  
 1255 AGGTACTCACCCACAGACAAATGCAATTTTAACTTTTAAAGAAAGACATCC 1314  
 1454 AGGTATTCACCCACAGACAAATGCAATTTTAACTTTTAAAGACCAAGACAGCC 1512  
 1315 AGTATCTGATCTTTTGAAGACCTGATTAATCAATCAAGCACTCT 1360  
 1513 AATAGTTAAGATCTTTTGAAGACCTGATTAATCAATCAAGCACTCT 1558

RESULT 15  
 156872 156872 1600 bp DNA 1linear PAT 07-OCT-1997  
 LOCUS Sequence 5 from patent US 5650501.  
 DEFINITION 156872  
 ACCESSION 156872  
 VERSION 156872.1 GI:2477285  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1600)  
 AUTHORS Dennis J.W., Hefner M., and Fode C.  
 TITLE Serine/threonine kinase and nucleic acid encoding same  
 JOURNAL Patent: US 5650501-A 5 22-JUL-1997;  
 FEATURES  
 source location/Qualifiers  
 1..1600  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 33.3%; Score 970.4; DB 6; Length 1600;  
 Best Local Similarity 85.0%; Pred. No. 1.1e-200;  
 Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

1 ATGGGCACTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTGTA 60  
 206 ATGGGGCGCTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTGTA 265  
 61 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 120  
 266 GGATCATTTGCTGTGTCTTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 325  
 121 AAAATGATGATPAAGAAAGCATGTACAAAGCAGAAAGTGTACAGAGTCCAAATGAG 180  
 326 AAAATGATGATPAAGAAAGCATGTACAAAGCAGAAAGTGTACAGAGTCCAAATGAG 385  
 181 GTGAAAATACATTTGCCAATTTGAAACATCCCTTATCTTGAGCTTTATTAATTTTGA 240  
 386 GTGAAAATACATTTGCCAATTTGAAACATCCCTTATCTTGAGCTTTATTAATTTTGA 445  
 241 GATAGCAATTAATGTATCTGTGATTAAGAAATGTCCATTAATGAGAAATGAACAGTAT 300  
 446 GATTAACAATTAATGTATCTGTGATTAAGAAATGTCCATTAATGAGAAATGAACAGTAT 505  
 301 CTAAAGAAATGAGTGAACCCCTTCTGAGAAATGAAGCTCGACATTCATGACCGAGATC 360  
 506 CTGAAGAAACAGATGAAGACCTTTCTGAGAAAGGAAAGCTTAGAGCTTCATGACCGAGAT 565

```
OY 361 ATCAGAGGATGTTGATCTTCAATTCAGTGGTACTACACCGGAGCCTCACACTTTCT 420
    |||||
Db 566 ATCAGAGGATGTTGATCTTCAATTCAGTGGTACTACACCGGAGCCTCACACTTTCT 625
OY 421 AACCTCTACTGATCTGTAATATGAACATCAAGATTGCTGATTTTGGGCTGGCACTGA 480
    |||||
Db 626 AACATCTTACTTACGGGAAATATGAACATTAATAATGCTGACTTGTGACTAGCAACGCA 685
OY 481 CTGAATAATGCCAATGAAGAAGCATATATATGTGAACTCCCTAACATTAATTCACCA 540
    |||||
Db 686 TTGAATATGCCAATGAAGAAGCATATATATGTGAACTCCCTAACATTAATTCACCA 745
OY 541 GAAATTTGCCACTGAAAGTCACATGGCCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600
    |||||
Db 746 GAAATTTGCCACTGAAAGTCACATGGCCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 805
OY 601 TTTTATACATTAATTATCGGAGACCACTTTGACACTGACACAGTCAAGAACACATTA 660
    |||||
Db 806 TCTTATACGTTACTTATTTGGAAGACCACTTTTGAACACTGACACAGTCAAGAACATTA 865
OY 661 AATAAGTGTATTTGGCAATTAATGAATGCCATCTTTTGTCAATAGAGGCCAAGAC 720
    |||||
Db 866 AACAAAGTAGTCTTGGCAGATTATGAATGCCACCTTTTGTCAAGAGGCCAAGAC 925
OY 721 CTTAATTCACAGTTACTTGTAGAAATCCAGCAGATCGTTTAACTGTCTTCAATATG 780
    |||||
Db 926 CTTAATTCACAGTTACTTGTAGAAATCCAGCAGATCGTTTAACTGTCTTCAATATG 985
OY 781 GACCATCTTTTATATGTCGGAAATCTTCAACAAAAGTAAAGTTAAGAACTGTGAA 840
    |||||
Db 986 GACCATCTTTTATATGTCGGAAATCTTCAACAAAAGTAAAGTAAAGTAAAGTAAAG 1045
OY 841 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAACGCTTCCAGTACC 900
    |||||
Db 1046 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAACGCTTCCAGTACC 1105
OY 901 AGTATATGATGATAGTTATTTTGAACAAAGAGACTTTGATTGTCAGCCTCCCAAT 960
    |||||
Db 1106 AGTATATGATGATAGTTATTTTGAACAAAGAGACTTTGATTGTCAGCCTCCCAAT 1162
OY 961 AAAATGCTGATATTTCCAAAGAAATTAAGTCACTGATTTTCTTCTCAGAGATGA 1020
    |||||
Db 1021 AAAATGCTGATATTTCCAAAGAAATTAAGTCACTGATTTTCTTCTCAGAGATGA 1074
OY 1075 AAGATATTCAGATGACAGAAAGAGGCAATTTCTGATACCTTGTAGAGCTTATTC 1134
    |||||
Db 1280 AAGATATTCAGATGACAGAAAGAGGCAATTTCTGATACCTTGTAGAGCTTATTC 1339
OY 1135 TCTGATATGATCTGGCACTTCTAATAGTCACTCAAGCAAAAACATATTAATGAGACA 1194
    |||||
Db 1340 TCTGATATGATCTGGCACTTCTAATAGTCACTCAAGCAAAAACATATTAATGAGACA 1396
OY 1195 TGTCACTCGAGCAAAATGCTTTCAGTGTCCAAAGATCAAGAGAGGTGAAAATGAAGAG 1254
    |||||
Db 1397 TGTCACTCGAGCAAAATGCTTTCAGTGTCCAAAGATCAAGAGAGGTGAAAATGAAGAG 1453
OY 1255 AAGTATCTCAACCCACAGAACAAATGCCAATTTTAACTTTTAAAGAAAAGACATCC 1314
    |||||
Db 1454 AAGTATCTCAACCCACAGAACAAATGCCAATTTTAACTTTTAAAGAAAAGACATCC 1512
OY 1315 AAGTATCTCGATCTTTTGAAGACCTGATTAACATCAAGCACTCT 1360
    |||||
Db 1513 AATAGTTAAGATCTTTTGAAGACCTGATTAATGACTGAGCAGTAT 1558
    |||||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 12:53:07 ; Search time 1076 Seconds  
(without alignments)  
11500.930 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913  
Sequence: 1 acggcgaccatgcacgsggga.....cgaccctatttcattga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: genebegn1980s:\*\n2: genebegn1980s:\*\n3: genebegn2000s:\*\n4: genebegn2001as:\*\n5: genebegn2001bs:\*\n6: genebegn2002s:\*\n7: genebegn2003as:\*\n8: genebegn2003bs:\*\n9: genebegn2003cs:\*\n10: genebegn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3937	4	AAK51966 Human pol
2	2911.4	99.9	3937	4	AAI58400 Human pol
3	2911.4	99.9	3937	8	ADB48370 Novel hum
4	2908.2	99.8	3721	8	ACH03903 Human CDN
5	2825.8	97.0	3841	4	AAK52950 Human pol
6	2825.8	97.0	3845	4	AAI60186 Human pol
7	1879	64.5	3447	6	ABI99713 Mouse lsc
8	1878	64.5	3447	2	AAI08711 Sak-a ser
9	970.4	33.3	1599	2	AAI08712 Sak-b ser
10	924.2	31.7	1453	2	AAI08710 Sak serin
11	426	14.6	484	8	ACH34348 Human end
12	329.2	11.3	609	6	AB136507 Human col
13	284	9.7	609	9	ADD33867 Mouse mit
14	282.2	9.7	2585	4	AB112501 Drosophi1
15	250	8.6	491	8	ACH37505 Human end
16	239	8.2	4961	4	AB112500 Drosophi1
17	216.8	7.4	302	7	ABT40590 Toxicity
18	154.2	5.3	381	7	ACA56051 Human sig
19	135	4.6	2772	6	ABI99219 Mouse lsc
20	134.4	4.6	2770	6	AAK06831 Disease a
21	134.4	4.6	2783	6	ABQ91962 Human NF-
22	134.4	4.6	2789	2	AAK27228 Human ser
23	134.4	4.6	2976	6	ABQ55069 Human ova

24	133.4	4.6	2781	7	ABT42251 Toxicity
25	133.4	4.6	2781	9	ADB53360 Primary r
26	133.2	4.6	2226	9	ADB59070 Toxicity-
27	133.2	4.6	2226	9	ADB53819 Primary r
28	132.4	4.5	204	4	AAK06737 Polynucle
29	132.4	4.5	2783	2	AAK27227 Human ser
30	131.4	4.5	2169	6	ABK84377 Human CDN
31	131.4	4.5	2169	7	ACA56549 Human sig
32	131.4	4.5	2169	7	ACA52787 Human PRK
33	131.4	4.5	2169	8	ADA50173 Human pol
34	130.2	4.5	2623	4	AB111407 Drosophi1
35	124	4.3	2778	5	AAH16619 Human CDN
36	124	4.3	2778	5	AAH78069 Nucleotid
37	123.4	4.2	2248	3	ACB83112 DNA encod
38	119	4.1	1485	3	ACA6066 Arabidops
39	118	4.1	1309	9	ACF58068 Human Hsa
40	118	4.1	2033	6	ABK64870 DNA encod
41	118	4.1	2033	9	ADD89968 Human can
42	115.2	4.0	1539	6	AB212357 Arabidops
43	115.2	4.0	1792	3	ACA44825 Arabidops
44	115.2	4.0	1825	6	AAI46639 A thalian
45	115.2	4.0	1825	6	AAI46642 A thalian

## ALIGNMENTS

RESULT 1	AAK51966	AAK51966 standard; cDNA, 3937 BP.
ID	AAK51966	
XX	AAK51966;	
AC	AAK51966;	
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human polynucleotide SEQ ID NO 511.	
XX		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200157190-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	05-FEB-2001; 2001WC-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI	Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZM;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
DR	WPI; 2001-476283/51.	
DR	P-PSDB; AAM78833.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
XX		
PS	Claim 1; Page 1846-1850; 6221pp; English.	
XX		

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW7323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX

Sequence 3937 BP; 1295 A; 732 C; 773 G; 1137 T; 0 U; 0 Other;

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCGACTGCAATCGGAGAGAGATGAGAGATTTTAAAGTTGGAAATCGCTTGGTAAA 60
Db 334 ATGGCGACTGCAATCGGAGAGAGATGAGAGATTTTAAAGTTGGAAATCGCTTGGTAAA 393
QY 61 GATCATTTTCTGGTGTCTACAGAGCTGAGTCCATTACACTGGTTTGGAAAGTTGCAATC 120
Db 394 GATCATTTTCTGGTGTCTACAGAGCTGAGTCCATTACACTGGTTTGGAAAGTTGCAATC 453
QY 121 AAATATGATATAGAAAGCCATGTACAAAGCAGAAATGTTACAGAGATCCAAAATGAG 180
Db 454 AAATATGATATAGAAAGCCATGTACAAAGCAGAAATGTTACAGAGATCCAAAATGAG 513
QY 181 GTGAAAATACATTGGCCAAATGAAACATCCTTATCTTGGAGCTTATATCTATTTGAA 240
Db 514 GTGAAAATACATTGGCCAAATGAAACATCCTTATCTTGGAGCTTATATCTATTTGAA 573
QY 241 GATAGCAATTATGTATCTGATATAGAAATGTGCATATATGAGAAATGAAACAGTAT 300
Db 574 GATAGCAATTATGTATCTGATATAGAAATGTGCATATATGAGAAATGAAACAGTAT 633
QY 301 CTAAAGATATAGATGAAACCTTCTCAGAAATATGAAAGTCGACATTCATGACAGATC 360
Db 634 CTAAAGATATAGATGAAACCTTCTCAGAAATATGAAAGTCGACATTCATGACAGATC 693
QY 361 ATCAGAGGATATGTATCTTCAATTCATGTATATCTACACCGGAGACTCAACTTTCT 420
Db 694 ATCAGAGGATATGTATCTTCAATTCATGTATATCTACACCGGAGACTCAACTTTCT 753
QY 421 AACCTCTCTGACTCGTATATATGAAACATCAAGATTGCGATTTGGGCTGGCACTCAA 480
Db 754 AACCTCTCTGACTCGTATATATGAAACATCAAGATTGCGATTTGGGCTGGCACTCAA 813
QY 481 CTGAAAATATGCCATGAAAAGCACTATACATTTATGTGAACTCTTAACTATATTCACA 540
Db 814 CTGAAAATATGCCATGAAAAGCACTATATATGTGAACTCTTAACTATATTCACA 873
QY 541 GAAATTTGCGACATCGAAGTGCATATGAGCTTTGATATGATTTGGTCCCTGGGCTGTATG 600
Db 874 GAAATTTGCGACATCGAAGTGCATATGAGCTTTGATATGATTTGGTCCCTGGGCTGTATG 933
QY 601 TTTTATACATTACTTATCGGAGAGCAACCTTTCGACACTGACAGTCAAGAACACATTA 660
Db 934 TTTTATACATTACTTATCGGAGAGCAACCTTTCGACACTGACAGTCAAGAACACATTA 993
QY 661 AATAAAGATATGTTGGCAGATTATGAAATGCAATCTTTTGTGCAATATAGAGCCCAAGAC 720
Db 994 AATAAAGATATGTTGGCAGATTATGAAATGCAATCTTTTGTGCAATATAGAGCCCAAGAC 1053
QY 721 CTATATTCACAGTACTCTGTATGAAATCAGCAGATGTTTAACTGTCTTCACTGATATG 780
Db 1054 CTATATTCACAGTACTCTGTATGAAATCAGCAGATGTTTAACTGTCTTCACTGATATG 1113
QY 781 GACCATCTTTTATATGTCGGAATTTCTTCAACAAAAGTAAAGATTTAGAACCTGTGAA 840
```

```
Db 1114 GACCATCTTTTATATGTCGGAATTTCTTCAACAAAAGTAAAGATTTAGAACCTGTGAA 1173
QY 841 GACTCAATGATATGAGGAGATGCCACATTTTCTACATTTACAGCTTTCTTCCAGTACC 900
Db 1174 GACTCAATGATATGAGGAGATGCCACATTTTCTACATTTACAGCTTTCTTCCAGTACC 1233
QY 901 AGTATAGTGTATGATTTATTTGACAAAAGAACATTTGATTGGTCAGCCATCCCAAT 960
Db 1234 AGTATAGTGTATGATTTATTTGACAAAAGAACATTTGATTGGTCAGCCATCCCAAT 1293
QY 961 AAAATGACATGATTTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAGAGATGGA 1020
Db 1294 AAAATGACATGATTTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAGAGATGGA 1353
QY 1021 AACAGTTTATATCTCAGTGGGGAATATAGAAACCAATATATGAGAAAGGGAAGATA 1080
Db 1354 AACAGTTTATATCTCAGTGGGGAATATAGAAACCAATATATGAGAAAGGGAAGATA 1413
QY 1081 ATTCAAGATGCAAGAAAGGCCAATTCCTGATACCTTCGTAGAGCTTATTCCTGTAT 1140
Db 1414 ATTCAAGATGCAAGAAAGGCCAATTCCTGATACCTTCGTAGAGCTTATTCCTGTAT 1473
QY 1141 AGATCTGGCACTTCTTAATATGTCAGTCTCAGCAAAAACATATATACATGAAACGATGCAC 1200
Db 1474 AGATCTGGCACTTCTTAATATGTCAGTCTCAGCAAAAACATATATACATGAAACGATGCAC 1533
QY 1201 TCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGGTGAAATGAAAGAGGTAC 1260
Db 1534 TCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGGTGAAATGAAAGAGGTAC 1593
QY 1261 TCACCCACAGCAACATATGCCAATTTTATCTTCTTAAAGAAAAGACATCCAGTAGT 1320
Db 1594 TCACCCACAGCAACATATGCCAATTTTATCTTCTTAAAGAAAAGACATCCAGTAGT 1653
QY 1321 TCTGATCTTTTGAAGAACCTGATATACATCAAGCACTTCCCAATCATCTTTGTCCAGGA 1380
Db 1654 TCTGATCTTTTGAAGAACCTGATATACATCAAGCACTTCCCAATCATCTTTGTCCAGGA 1713
QY 1381 AAAAATCTCTTTCATTTGCAATTTGCAACCCGACCTCAGACTGAAACCGTACAAAGTGGTT 1440
Db 1714 AAAAATCTCTTTCATTTGCAATTTGCAACCCGACCTCAGACTGAAACCGTACAAAGTGGTT 1773
QY 1441 GGGATCTGCAAAATTAATGCTCATTTTAAGAAAACATCTGAATATATGACAGATCAAGCCCA 1500
Db 1774 GGGATCTGCAAAATTAATGCTCATTTTAAGAAAACATCTGAATATATGACAGATCAAGCCCA 1833
QY 1501 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACACATCAAAAATATGCTGAGACT 1560
Db 1834 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACACATCAAAAATATGCTGAGACT 1893
QY 1561 GATACAAAAGTCAAAAAGAACTGTATGCTTCTGATATATGACATTTCTGTAACACGCAA 1620
Db 1894 GATACAAAAGTCAAAAAGAACTGTATGCTTCTGATATATGACATTTCTGTAACACGCAA 1953
QY 1621 AATACATGAAATATATATGACTGCACTTCAAGTAACTGAGTAACTCAACAAAGATAT 1680
Db 1954 AATACATGAAATATATATGACTGCACTTCAAGTAACTGAGTAACTCAACAAAGATAT 2013
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAAAGCTAGGGGATATGAGCCACCATAG 1740
Db 2014 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAAAGCTAGGGGATATGAGCCACCATAG 2073
QY 1741 GGTATACAAATGCTATTAAGAGCATTAATCTCCCTGGTGGTCTCAAGGTTAAAA 1800
Db 2074 GGTATACAAATGCTATTAAGAGCATTAATCTCCCTGGTGGTCTCAAGGTTAAAA 2133
QY 1801 CCAATCAGACAGAAAACCAAAAAGCTGTGTAGACATCTTGAATTCAGAGAGAGGTGTGT 1860
Db 2134 CCAATCAGACAGAAAACCAAAAAGCTGTGTAGACATCTTGAATTCAGAGAGAGGTGTGT 2193
QY 1861 GTGAGCTTGTAAAGAGATATGATCTCAAGAAATATGAAAGAAAGTTCTTCAATATCT 1920
```

```

Db 2194 GTGAGCTGTTGAAAGAGTATGATCTCAAGATATATGTAAGAGTTCTTCAGATATCT 2253
Qy 1921 AGTATGAAATATGATCACTATTTATTTATCCAAATGATGATGAGTTTCTCTGCT 1980
Db 2254 AGTATGAAATATGATCACTATTTATTTATCCAAATGATGATGAGTTTCTCTGCT 2313
Qy 1981 GATAGACACCCCTCACTACTGACACATCAGTAGGTACAGCTTTGACATTTACAGAA 2040
Db 2314 GATAGACACCCCTCACTACTGACACATCAGTAGGTACAGCTTTGACATTTACAGAA 2373
Qy 2041 AATATCTGCGAAATATCATATGCTTCCAGTTTGTACAGCTTTGAAGATCTAAATCT 2100
Db 2374 AATATCTGCGAAATATCATATGCTTCCAGTTTGTACAGCTTTGAAGATCTAAATCT 2433
Qy 2101 CCCAAATATCATATTTTCAAGATATGCTTAAATGATTTGATGAGAAATCTCTGCT 2160
Db 2434 CCCAAATATCATATTTTCAAGATATGCTTAAATGATTTGATGAGAAATCTCTGCT 2493
Qy 2161 GGTGATTTTGAAGTTTGGTTTATGATGAGGTTAAATACACAAACAGAAATTCATT 2220
Db 2494 GGTGATTTTGAAGTTTGGTTTATGATGAGGTTAAATACACAAACAGAAATTCATT 2553
Qy 2221 CAGGTGATTTGAAAGACAGGGAAGTCTTACATTTAAAAAGTGAAGTAAATAC 2280
Db 2554 CAGGTGATTTGAAAGACAGGGAAGTCTTACATTTAAAAAGTGAAGTAAATAC 2613
Qy 2281 TTGAAAGAGAGATTAATAATGATATGATGACATGCTTAATGAGGTTATGATTTGTTA 2340
Db 2614 TTGAAAGAGAGATTAATAATGATATGATGACATGCTTAATGAGGTTATGATTTGTTA 2673
Qy 2341 GCACTGGAATCCATATTTTCAAGAGAGAAAGGAAATAGAGAGTCTCTCTTTTCCCA 2400
Db 2674 GCACTGGAATCCATATTTTCAAGAGAGAAAGGAAATAGAGAGTCTCTCTTTTCCCA 2733
Qy 2401 ATATATATGAAAGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2734 ATATATATGAAAGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2793
Qy 2461 TCTGTGATTTCAATTAATTAACCAACGAGATGATGATGATGATGATGATGATGATGAT 2520
Db 2794 TCTGTGATTTCAATTAATTAACCAACGAGATGATGATGATGATGATGATGATGATGAT 2853
Qy 2521 AGTGTGCTTCTCCACACAGGACCAATCTTAATCTCTTAATGATGATGATGATGATGAT 2580
Db 2854 AGTGTGCTTCTCCACACAGGACCAATCTTAATCTCTTAATGATGATGATGATGATGAT 2913
Qy 2581 CTTGTGCTTCAACATCAAGCTTCTGGAACAGATCTCTTAATGATGATGATGATGATGAT 2640
Db 2914 CTTGTGCTTCAACATCAAGCTTCTGGAACAGATCTCTTAATGATGATGATGATGATGAT 2973
Qy 2641 CTTCTTAATCAAGACATCTTTGAAATCTGTTTGTGAAATGTTGTTGTTGTTGTTGTTA 2700
Db 2974 CTTCTTAATCAAGACATCTTTGAAATCTGTTTGTGAAATGTTGTTGTTGTTGTTA 3033
Qy 2701 CAGTTTACATGAGGAGCTGTTGAGTTTCAAGTTTAAATGATGAGGTTGTTGTTGTTG 2760
Db 3034 CAGTTTACATGAGGAGCTGTTGAGTTTCAAGTTTAAATGATGAGGTTGTTGTTGTTG 3093
Qy 2761 GGAGAGTGTCTTCTATCACTATTAATCTCAACCAATGTTCAACATCAAGTATGAGAA 2820
Db 3094 GGAGAGTGTCTTCTATCACTATTAATCTCAACCAATGTTCAACATCAAGTATGAGAA 3153
Qy 2821 AATGAAATATCAAGATCACTATTAATCTCAACCAATGTTCAACATCAAGTATGAGAA 2880
Db 3154 AATGAAATATCAAGATCACTATTAATCTCAACCAATGTTCAACATCAAGTATGAGAA 3213
Qy 2881 ATGTTTCTATCACTGCTCTATTTTCAATGA 2913
Db 3214 ATGTTTCTATCACTGCTCTATTTTCAATGA 3246

```

RESULT 2  
AA158400

```

ID AA158400 standard; cDNA; 3937 BP.
XX
AC AA158400;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 603.
XX
KW Human; nocrotic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HSP-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dermanac RT;
PI
XX
DR WPI, 2001-442253/47.
DR F-PSDB; AAM39244.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 603; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nocrotic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 3937 BP; 1295 A; 732 C; 773 G; 1137 T; 0 U; 0 Other;

```

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 ATGGGCACTGATGGGGAAGATCGAGATTTTAAAGTTGAAATCTGTTGGTAA 60
Db 334 ATGGGCACTGATGGGGAAGATCGAGATTTTAAAGTTGAAATCTGTTGGTAA 393

```

QY 61 GAATCATTTGCTGTGTCTACAGAGCTGAGTCCATTCACTGTTTGGAAAGTTGCAGTC 120  
Db 394 GGATCATTTTGTCTGTGTCTACAGAGCTGAGTCCATTCACTGTTTGGAAAGTTGCAGTC 453  
QY 121 AAAATGATGATTAAGAAAGCCATGTACAAAGCCGAATGTTACAGAGAGTCCAAAATGAG 180  
Db 454 AAAATGATGATTAAGAAAGCCATGTACAAAGCCGAATGTTACAGAGAGTCCAAAATGAG 513  
QY 181 GTGAAAATACATTTGCCAATTGAAACATCTTCTATCTTTGGAGCTTTATACTATTTTGA 240  
Db 514 GTGAAAATACATTTGCCAATTGAAACATCTTCTATCTTTGGAGCTTTATACTATTTTGA 573  
QY 241 GATAGCAATTATGTATCTGTATTAAGAAATGTGCCATATAGAGAAAATGAAACAGTAT 300  
Db 574 GATAGCAATTATGTATCTGTATTAAGAAATGTGCCATATAGAGAAAATGAAACAGTAT 633  
QY 301 CTAAAGATTAAGTGAACCCCTTCTGAGAAAATGAAGCTTGACACTTTCATGACCGAGTC 360  
Db 634 CTAAAGATTAAGTGAACCCCTTCTGAGAAAATGAAGCTTGACACTTTCATGACCGAGTC 693  
QY 361 ATCACAGGGATGTTGTATCTTCAATGTAATCTACACCGGAGCCCTCACATTTCT 420  
Db 694 ATCACAGGGATGTTGTATCTTCAATGTAATCTACACCGGAGCCCTCACATTTCT 753  
QY 421 AACCTCTTCTGACTCTGTATATGAACATCAAGATTTGAGTTGGGCTGGCACTCAA 480  
Db 754 AACCTCTTCTGACTCTGTATATGAACATCAAGATTTGAGTTGGGCTGGCACTCAA 813  
QY 481 CTGAAAATGCCACATGAAGGACCTATACATTAATGAGAACTCCCTAACCTAACCTTTCACA 540  
Db 814 CTGAAAATGCCACATGAAGGACCTATACATTAATGAGAACTCCCTAACCTAACCTTTCACA 873  
QY 541 GAAATTTGCCACTGGAAGTGACATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600  
Db 874 GAAATTTGCCACTGGAAGTGACATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 933  
QY 601 TTTTATACATTAATCTATCGGAGACCAAGCTTGTGACATGACACAGTCAAGAACATTA 660  
Db 934 TTTTATACATTAATCTATCGGAGACCAAGCTTGTGACATGACACAGTCAAGAACATTA 993  
QY 661 AATTAAGTATGATGAGCAATTAATGAATGCCATCTTTTGTCAATAGAGGCAAGGAC 720  
Db 994 AATTAAGTATGATGAGCAATTAATGAATGCCATCTTTTGTCAATAGAGGCAAGGAC 1053  
QY 721 CTATATCACAGTTACTTGTGAAATCCAGACAGTCTGTTAAGTCTGTCTTCAATATG 780  
Db 1054 CTATATCACAGTTACTTGTGAAATCCAGACAGTCTGTTAAGTCTGTCTTCAATATG 1113  
QY 781 GACCATCTTTTATGTCCTCCGAAATCTTCAACAAAAGTAAAGATTAGAGAACTGTGAA 840  
Db 1114 GACCATCTTTTATGTCCTCCGAAATCTTCAACAAAAGTAAAGATTAGAGAACTGTGAA 1173  
QY 841 GACTCAATTGATAGTGGGATGCCCAATTTCTACTGCAATTAAGCTTCTTCCAGTACC 900  
Db 1174 GACTCAATTGATAGTGGGATGCCCAATTTCTACTGCAATTAAGCTTCTTCCAGTACC 1233  
QY 901 AGTATTAAGTGTAGTTATTTTGAACAAAGAGACTTTGATTTGTCAGGCCATCTCCAAAT 960  
Db 1234 AGTATTAAGTGTAGTTATTTTGAACAAAGAGACTTTGATTTGTCAGGCCATCTCCAAAT 1293  
QY 961 AAAATGACGTATTTCCAAAGAAATAAGTTCAACTGATTTTTCTTCTTCAGAGATGGA 1020  
Db 1294 AAAATGACGTATTTCCAAAGAAATAAGTTCAACTGATTTTTCTTCTTCAGAGATGGA 1353  
QY 1021 AACAGTTTTTATATCTCAGTGGGGAATCAAGAAACAGTAATAGTGAAGGGAGAGATA 1080  
Db 1354 AACAGTTTTTATATCTCAGTGGGGAATCAAGAAACAGTAATAGTGAAGGGAGAGATA 1413  
QY 1081 ATTCAAGTGCAGAGAAAGGCCCAATTCCTGATACCTTTCGTAAGCTTATTTCTCTGAT 1140  
Db 1414 ATTCAAGTGCAGAGAAAGGCCCAATTCCTGATACCTTTCGTAAGCTTATTTCTCTGAT 1473  
QY 1414 AGATGTGCACTTCTTAATATGTAGTCTCAGAGAAAACATATACATAGGAACGATGTAC 1200

Db 1474 AGATGTGCACTTCTTAATATGTAGTCTCAGAGAAAACATATACATAGGAACGATGTAC 1533  
QY 1201 TCAGCAAGAAATCTTTCAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTAC 1260  
Db 1534 TCAGCAAGAAATCTTTCAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTAC 1593  
QY 1261 TCACCAACAGACAAAGAGCCAACTTTTAACTTCTTAAAGAAAGACATCCGATGT 1320  
Db 1594 TCACCAACAGACAAAGAGCCAACTTTTAACTTCTTAAAGAAAGACATCCGATGT 1653  
QY 1321 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGACCTCTCAATCATCTTTGTCCAGA 1380  
Db 1654 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGACCTCTCAATCATCTTTGTCCAGA 1713  
QY 1381 AAAATCTCTTTTCCATTTGACAGCCGACACTCGACTGAACCGTACAAAGTGT 1440  
Db 1714 AAAATCTCTTTTCCATTTGACAGCCGACACTCGACTGAACCGTACAAAGTGT 1773  
QY 1441 GGGATCTGCAAAATTAATGCTCATTTTAAAGAAAACCTAGTAATAGACAGATCAGCCCA 1500  
Db 1774 GGGATCTGCAAAATTAATGCTCATTTTAAAGAAAACCTAGTAATAGACAGATCAGCCCA 1833  
QY 1501 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAATGCTGAGCT 1560  
Db 1834 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAATGCTGAGCT 1893  
QY 1561 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGTAATGCACTTCTGTAAAACAGAA 1620  
Db 1894 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGTAATGCACTTCTGTAAAACAGAA 1953  
QY 1621 AATACATGAATATATGACTGCACTTCAACAGTAAACCTGAGATATCCAAAGAAATGT 1680  
Db 1954 AATACATGAATATATGACTGCACTTCAACAGTAAACCTGAGATATCCAAAGAAATGT 2013  
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAAAGCTAGGGGTATGAGCCCAATGG 1740  
Db 2014 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAAAGCTAGGGGTATGAGCCCAATGG 2073  
QY 1741 GGTATACGAATCGTACATTAAGAAAGATTAACATCTCGTTGGTGTCTCACAGTTTAAA 1800  
Db 2074 GGTATACGAATCGTACATTAAGAAAGATTAACATCTCGTTGGTGTCTCACAGTTTAAA 2133  
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCATCTTGAATCAGAGAGTGTGT 1860  
Db 2134 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCATCTTGAATCAGAGAGTGTGT 2193  
QY 1861 GTGAGCTTGTAAAGAGATGATCTTCAAGAAATATGTGAAGAAAGTTCTTCAATATCT 1920  
Db 2194 GTGAGCTTGTAAAGAGATGATCTTCAAGAAATATGTGAAGAAAGTTCTTCAATATCT 2253  
QY 1921 AGTATGGAATATGATCACTATTTATTCCAAATGAGTGAAGGTTTCCCTTGGCT 1980  
Db 2254 AGTATGGAATATGATCACTATTTATTCCAAATGAGTGAAGGTTTCCCTTGGCT 2313  
QY 1981 GATAGCACACCTTCACTGACACACATGATGATAGTATGACGTTTGAACAATTTACAGAA 2040  
Db 2314 GATAGCACACCTTCACTGACACACATGATGATAGTATGACGTTTGAACAATTTACAGAA 2373  
QY 2041 AAAATCTGGGGAATAATCAATATGCTTCCAGGTTTGTACGCTTGTAAATCTAAATCT 2100  
Db 2374 AAAATCTGGGGAATAATCAATATGCTTCCAGGTTTGTACGCTTGTAAATCTAAATCT 2433  
QY 2101 CCCAAATCACTTATTTTCAAGATATGCTAAATGCAATTTGATGGAATTCCTCGGT 2160  
Db 2434 CCCAAATCACTTATTTTCAAGATATGCTAAATGCAATTTGATGGAATTCCTCGGT 2493  
QY 2161 GCTGATTTGAGGTTGGTTTATGATGGGTAAATAACAACAAACAGAGATTTCAAT 2220  
Db 2494 GCTGATTTGAGGTTGGTTTATGATGGGTAAATAACAACAAACAGAGATTTCAAT 2553  
QY 2221 CAGGTGATTTGAAAAGACAGGGAAGTCTTACCTTTAAAAAGTGAAGTTAATAGC 2280

Db	2551	CAGGTGATGTAAAAAGACAGGGAGCTTTACACTTTAAAAAGTGAAGTAATATAGC	2613
Qy	2281	TTGAAAGAGAGATPAAAAATGTATATGAGCACTGCTAAATGAGGTCATCGTATTTGTTTA	2340
Db	2614	TTGAAAGAGGAGTAAAAATGTATATATGAGCACTGCTAATGAGGTCATCGTATTTGTTTA	2673
Qy	2341	GCACTGGAAATCCATTAATTTTCAGAAGAGAAAGAAAACTAGAGAGTCCCTTTTCCCA	2400
Db	2674	GCACTGGAAATCCATTAATTTTCAGAAGAGAAAGAAAACTAGAGAGTCCCTTTTCCCA	2733
Qy	2401	ATAATCATAGAGAGAAAACTTGTAGTACTATGTTCACTTAAGGCTTATCACTCTCT	2460
Db	2734	ATAATCATAGAGAGAAAACTTGTAGTACTATGTTCACTTAAGGCTTATCACTCTCT	2793
Qy	2461	TCCTGTGATTCAAATTAATCCCAACAGAGATAGACATCTTTCAACAGAAATGTCATGAT	2520
Db	2794	TCCTGTGATTCAAATTAATCCCAACAGAGATAGACATCTTTCAACAGAAATGTCATGAT	2853
Qy	2521	AGTCTGCTCTTCCCAACAGAGCAACAATCCTTAATCCCTCTATGTGTTACAAATGAAGA	2580
Db	2854	AGTCTGCTCTTCCCAACAGAGCAACAATCCTTAATCCCTCTATGTGTTACAAATGAAGA	2913
Qy	2581	CTTGTGCTTACACTACAGCTTCTGGAACAGACATCTCTTCTAATAGTCTAAAAAGTTGT	2640
Db	2914	CTTGTGCTTACACTACAGCTTCTGGAACAGACATCTCTTCTAATAGTCTAAAAAGTTGT	2973
Qy	2641	CTTCTCTAAATCAGGACAACTTTTGAATCTGTTTTTGAAAAAATGTGTGGTGGCTACA	2700
Db	2974	CTTCTCTAAATCAGGACAACTTTTGAATCTGTTTTTGAAAAAATGTGTGGTGGCTACA	3033
Qy	2701	CAGTTAACTAGTGAAGCTGTGTGGGTTCACTTAATGAATGATGGTCCCAAGTTGTGTGCAG	2760
Db	3034	CAGTTAACTAGTGAAGCTGTGTGGGTTCACTTAATGAATGATGGTCCCAAGTTGTGTGCAG	3093
Qy	2761	GCAAGAGAGTCTTCTATCAGTTAATCCACCAATGCTCAACCACTAGGTATGAGGAA	2820
Db	3094	GCAAGAGAGTCTTCTATCAGTTAATCCACCAATGCTCAACCACTAGGTATGAGGAA	3153
Qy	2821	AATGAAAAAATTACAGACTACATCAAAACAGAAATTAACAGTGTCTGTCTTCATCTTTTG	2880
Db	3154	AATGAAAAAATTACAGACTACATCAAAACAGAAATTAACAGTGTCTGTCTTCATCTTTTG	3213
Qy	2881	ATGTTTTCTAATCCGACTCTTAATTTTCATTGA	2913
Db	3214	ATGTTTTCTAATCCGACTCTTAATTTTCATTGA	3246

XX	RESULT 3
ADBA48370	
ID	ADBA48370 standard; cDNA; 3937 BP.
XX	
AC	ADBA48370;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	Novel human cDNA SEQ ID NO 280.
XX	
KW	ss; cancer; neurodegenerative disease; human.
XX	
OS	Homo sapiens.
XX	
PN	US2003104529-A1.
XX	
PD	05-JUN-2003.
XX	
PF	04-JAN-2002; 2002US-00037270.
XX	
PR	21-JAN-2000; 2000US-00488725.
PR	25-APR-2000; 2000US-00552317.
PR	19-JUL-2000; 2000US-00620312.
XX	
PA	(ZHOU/) ZHOU P.
PA	(TANG/) TANG Y T.

Query Match	99.9%: Score 2911.4; DB 8; Length 3937;	Best Local Similarity 100.0%; Pred. No. 0;	Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	1 ATGCGACCTGCTCATCGGGAGAGATCGAGAGTTTAAAGTTGGAATCTGCTGTGTA	60	
DB	334 ATGCGACCTGCTCATCGGGAGAGATCGAGAGTTTAAAGTTGGAATCTGCTGTGTA	393	
QY	61 GGATCATTTGCTGTGCTCTACAGAGCTGATTCATTCACCTGTTTGGAACTGCATC	120	
DB	394 GGATCATTTGCTGTGCTCTACAGAGCTGATTCATTCACCTGTTTGGAACTGCATC	453	
QY	121 AAAATGATGATTAAGAAAGCATATGACAAGCAGAAATGATGACAGAGTCCAAATAG	180	
DB	454 AAAATGATGATTAAGAAAGCATATGACAAGCAGAAATGATGACAGAGTCCAAATAG	513	
QY	181 GTGAAATATATACATTTGCAATGTAATCATCTCTATCTTGAGACTTATATATATTTGAA	240	
DB	514 GTGAAATATATACATTTGCAATGTAATCATCTCTATCTTGAGACTTATATATATTTGAA	573	
QY	241 GATAGCAATTAATGTGTATCTGTATTTAGAAATGTGCCATAATGAGAAATGACAGTAT	300	
DB	574 GATAGCAATTAATGTGTATCTGTATTTAGAAATGTGCCATAATGAGAAATGACAGTAT	633	
QY	301 CTAAAGATTAAGTGAAGAAACCTCTTCAGAAATGAAGTCCAGCACTTATGACACAGATC	360	
DB	634 CTAAAGATTAAGTGAAGAAACCTCTTCAGAAATGAAGTCCAGCACTTATGACACAGATC	693	
QY	361 ATCAGAGGATGTGTATCTTCAATCTCATGTGTATATACACCGGACCTGACACTTCT	420	
DB	694 ATCAGAGGATGTGTATCTTCAATCTCATGTGTATATACACCGGACCTGACACTTCT	753	
QY	421 AACCTCTACTGACTCTGTATATATGAAACATCAAGATTCGTGATTTTGGCTGGCACTGAA	480	
DB	754 AACCTCTACTGACTCTGTATATATGAAACATCAAGATTCGTGATTTTGGCTGGCACTGAA	813	
QY	481 CTGAAATATGCACATGAAAGACCTATATCATTTATGTGGAATCTCTATATCACTTTACCA	540	
DB	814 CTGAAATATGCACATGAAAGACCTATATCATTTATGTGGAATCTCTATATCACTTTACCA	873	
QY	541 GAAATTCGACTCGAAGTGAATGAGCCTTGAATCTGATTTTGGTCCCTGGGCTGTATG	600	
DB	874 GAAATTCGACTCGAAGTGAATGAGCCTTGAATCTGATTTTGGTCCCTGGGCTGTATG	933	
QY	601 TTTTATATCAATTAATATGCGAGACCACTTCGACACTGACACAGTCAAGAACATTTA	660	
DB	934 TTTTATATCAATTAATATGCGAGACCACTTCGACACTGACACAGTCAAGAACATTTA	993	
QY	661 AATAAGTATGATTTGGCAGATTAATGAATGCACTTTTGTGTCAATAGAGCCCAAGAC	720	

D	b	994	AAATAAGATGATGGAGATTAGAAATGCAACTTTTGTCAATAGAGCCAAAGAC	1053
O	y	721	CTTATTCACCGATTACTTCGTAGAAATCAGACAGATGTTTAAGTGTCTTCAGATATG	780
D	b	1054	CTTATTCACCGATTACTTCGTAGAAATCAGACAGATGTTTAAGTGTCTTCAGATATG	1113
O	y	781	GACCATCTTTTATGATCCCGAAATTTCTTCAACAAAGTAAGATTAGAACTGTGAA	840
D	b	1114	GACCATCTTTTATGATCCCGAAATTTCTTCAACAAAGTAAGATTAGAACTGTGAA	1173
O	y	841	GACCATCTTTTATGATCCCGAAATTTCTTCAACAAAGTAAGATTAGAACTGTGAA	900
D	b	1174	GACCATCTTTTATGATCCCGAAATTTCTTCAACAAAGTAAGATTAGAACTGTGAA	1233
O	y	901	AGTAAAGTGTATGATTTATTTGACAAAGAAAGCTTTTGTGATGTCAGCCATCCCAAT	960
D	b	1234	AGTAAAGTGTATGATTTATTTTGTACAAAGAAAGAACTTTTGTGATGTCAGCCATCCCAAT	1293
O	y	961	AAATAGCTGTATTTCCAAAGAAATAAAGTTCAACTGATTTTCTTTCAGAGATGGA	1020
D	b	1294	AAATAGCTGTATTTCCAAAGAAATAAAGTTCAACTGATTTTCTTTCAGAGATGGA	1353
O	y	1021	AAAGTTTATATCTGATGGGGAATCAAGAAACCAATATATGTGAAAGGAGAGTA	1080
D	b	1354	AAAGTTTATATCTGATGGGGAATCAAGAAACCAATATATGTGAAAGGAGAGTA	1413
O	y	1081	ATTCAAGATGCAAGAAAGGAGCCATTTCTGATACCTTCGTAGACTTATTCCTGTAT	1140
D	b	1414	ATTCAAGATGCAAGAAAGGAGCCATTTCTGATACCTTCGTAGACTTATTCCTGTAT	1473
O	y	1141	AGATCTGGCATCTCTAATATGCTCACTGTCAGAGAAACATATACAAATGAAAGATGTAC	1200
D	b	1474	AGATCTGGCATCTCTAATATGCTCACTGTCAGAGAAACATATACAAATGAAAGATGTAC	1533
O	y	1201	TCAGCAAAATGCTTCACTGTCAGAGAAAGTACAGAGAGGTGAAATGAAAGAGTATC	1260
D	b	1534	TCAGCAAAATGCTTCACTGTCAGAGAAAGTACAGAGAGGTGAAATGAAAGAGTATC	1593
O	y	1261	TCACCAAGACAAATGCTCAATTTTAACTTTTAAAGAAAGATCTCAGTATG	1320
D	b	1594	TCACCAAGACAAATGCTCAATTTTAACTTTTAAAGAAAGATCTCAGTATG	1653
O	y	1321	TCGATCTTTTGAAGACCTGATTAACAATCAAGATCTTCAATCATCTTTTCCAGGA	1380
D	b	1654	TCGATCTTTTGAAGACCTGATTAACAATCAAGATCTTCAATCATCTTTTCCAGGA	1713
O	y	1381	AAAACTCTTTTCAATTTGACAGCCGACCTCAGACTGAACCGTCAACAGATGTTT	1440
D	b	1714	AAAACTCTTTTCAATTTGACAGCCGACCTCAGACTGAACCGTCAACAGATGTTT	1773
O	y	1441	GGGAATCTGCAAAATTAATGCTCAATTTAAGAAAACTAGTAATAGACATCAGCCCA	1500
D	b	1774	GGGAATCTGCAAAATTAATGCTCAATTTAAGAAAACTAGTAATAGACATCAGCCCA	1833
O	y	1501	AACCGGACCTTCAGGGCCATCCAGATTTGACAGAGACATCAAAAAAATGCTGACCT	1560
D	b	1834	AACCGGACCTTCAGGGCCATCCAGATTTGACAGAGACATCAAAAAAATGCTGACCT	1893
O	y	1561	GATACAAAGTCAAAAGAACTCTGATGCTTCTGATTAATGCAATCTGTAAAAACAGAA	1620
D	b	1894	GATACAAAGTCAAAAGAACTCTGATGCTTCTGATTAATGCAATCTGTAAAAACAGAA	1953
O	y	1621	AATACCAATGAAATATATGATCTGACCTTCAAGTAACTGTGATATCCAAAGAAATGT	1680
D	b	1954	AATACCAATGAAATATATGATCTGACCTTCAAGTAACTGTGATATCCAAAGAAATGT	2013
O	y	1681	GTTTTTGGCTCAGATCTCTTTTGAACAGAGCAAGACTAGGGATGAGGCGACATGG	1740
D	b	2014	GTTTTTGGCTCAGATCTCTTTTGAACAGAGCAAGACTAGGGATGAGGCGACATGG	2073
O	y	1741	GTTTATGAGATGTAATCTTAAGAAAGATTAATCTCGTTGTTGCTCAAGGTTAAAA	1800
D	b	2074	GTTTATGAGATGTAATCTTAAGAAAGATTAATCTCGTTGTTGCTCAAGGTTAAAA	2133
O	y	1801	CCAAATCAGACAGAAAAACAAAAAGCTGTGGAGACATCTGATTCAGAGAGGTGCT	1860
D	b	2134	CCAAATCAGACAGAAAAACAAAAAGCTGTGGAGACATCTGATTCAGAGAGGTGCT	2193
O	y	1861	GTTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAAGAAAGTTCTTCAATATCT	1920
D	b	2194	GTTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAAGAAAGTTCTTCAATATCT	2253
O	y	1921	AGTATGAGAAATAGATACATATTAATATCCAAATGTTGTGTAGAGTTTCTCTGCT	1980
D	b	2254	AGTATGAGAAATAGATACATATTAATATCCAAATGTTGTGTAGAGTTTCTCTGCT	2313
O	y	1981	GATAGACACCTCTCACTAGACAAATCATGATGATACAGCTTTGACATTTTACAGAA	2040
D	b	2314	GATAGACACCTCTCACTAGACAAATCATGATGATACAGCTTTGACATTTTACAGAA	2373
O	y	2041	AAATATGAGGAGAAATATCAATATGCTTCCAGTTTGTACAGCTTTGATGATTAATCT	2100
D	b	2374	AAATATGAGGAGAAATATCAATATGCTTCCAGTTTGTACAGCTTTGATGATTAATCT	2433
O	y	2101	CCCAAAATCACTTATTTTACAGATATGCTTAAATGCAATTTGATGAGAAATTCCTGCT	2160
D	b	2434	CCCAAAATCACTTATTTTACAGATATGCTTAAATGCAATTTGATGAGAAATTCCTGCT	2493
O	y	2161	GCTGATTTTGAAGTTTGTGTTTATGATGGGTAAATAACAAACAGAAAGATTTCTT	2220
D	b	2494	GCTGATTTTGAAGTTTGTGTTTATGATGGGTAAATAACAAACAGAAAGATTTCTT	2553
O	y	2221	CAGGTGATTTGAAAGACAGGGAGTCTTACCTTAAAGTAAAGTAAATGAAATATAGC	2280
D	b	2554	CAGGTGATTTGAAAGACAGGGAGTCTTACCTTAAAGTAAAGTAAATGAAATATAGC	2613
O	y	2281	TTGAAAGAGATTAATAATGATATATGACCATGCTAATAGGGTCACTGATTTGTTTA	2340
D	b	2614	TTGAAAGAGATTAATAATGATATATGACCATGCTAATAGGGTCACTGATTTGTTTA	2673
O	y	2341	GCACTGGAATCCATTAATTTGAGAAAGAAAGAAAGAAAGTGTCTCTTTTCCCA	2400
D	b	2674	GCACTGGAATCCATTAATTTGAGAAAGAAAGAAAGAAAGTGTCTCTTTTCCCA	2733
O	y	2401	ATAATCAATGAGAAAGAAAGCTGTGATGATCTGATGATGATGATGATGATGATGAT	2460
D	b	2734	ATAATCAATGAGAAAGAAAGCTGTGATGATCTGATGATGATGATGATGATGATGAT	2793
O	y	2461	TCGTGATTCATTAATCCCAAGAGATAGAGCATCTTTCAACAGAAATGTCATGAT	2520
D	b	2794	TCGTGATTCATTAATCCCAAGAGATAGAGCATCTTTCAACAGAAATGTCATGAT	2853
O	y	2521	AGTGTGCTCTTCCCAACAGGACCAATCCCTTAATGCTTACAAATGAAAGGA	2580
D	b	2854	AGTGTGCTCTTCCCAACAGGACCAATCCCTTAATGCTTACAAATGAAAGGA	2913
O	y	2581	CTTGTGCTTACATCAAGCTTCTGGAACAGATCTCTTCAATAGTCTTAAAGATTTGT	2640
D	b	2914	CTTGTGCTTACATCAAGCTTCTGGAACAGATCTCTTCAATAGTCTTAAAGATTTGT	2973
O	y	2641	CTTCTTAATCCAGCAAACTTTTGAATCTGTTTTTGTGAAAAATGTTGGTGGCTACA	2700
D	b	2974	CTTCTTAATCCAGCAAACTTTTGAATCTGTTTTTGTGAAAAATGTTGGTGGCTACA	3033
O	y	2701	CAGTTAATAGTGAAGCTGTGTGGTTCAAGTTAAATGATGCTCCAGTGTGTGAGAG	2760
D	b	3034	CAGTTAATAGTGAAGCTGTGTGGTTCAAGTTAAATGATGCTCCAGTGTGTGAGAG	3093
O	y	2761	GCAAGAGTGTCTTTTATGATGATTAATCTTCAACAAATGCTCAAACTAGTATGAGAA	2820
D	b	3094	GCAAGAGTGTCTTTTATGATGATTAATCTTCAACAAATGCTCAAACTAGTATGAGAA	3153
O	y	2821	AATGAAAAATTTACAGACATCAATCAAGAAATTAACAGTGTCTCTTCAATCTTTTG	2880
D	b	3154	AATGAAAAATTTACAGACATCAATCAAGAAATTAACAGTGTCTCTTCAATCTTTTG	3213



```
QY 1381 AAAATCCTTTTCATTTGACAGACCCGACACCTCAGACTGAAAACCGTTACAACAGTGGTTT 1440
    |||
Db 1668 AAAATCCTTTTCATTTGACAGACCCGACACCTCAGACTGAAAACCGTTACAACAGTGGTTT 1727
QY 1441 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTA CTGAATATGACAGACATCAGCCCA 1500
    |||
Db 1728 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTA CTGAATATGACAGACATCAGCCCA 1787
QY 1501 AACCGGAGCTTCCAGGGCCATCCAGATTTGAGAGAGACAATCAAAAATATGCTTGACAT 1560
    |||
Db 1788 AACCGGAGCTTCCAGGGCCATCCAGATTTGAGAGAGACAATCAAAAATATGCTTGACAT 1847
QY 1561 GATTCAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATCTCTGTAATAACAGCAA 1620
    |||
Db 1848 GATTCAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATCTCTGTAATAACAGCAA 1907
QY 1621 AATACCAATGAATATATGATGCTGACATTCAGATTAACCTGAGATTAATCCACAAGAAATGT 1680
    |||
Db 1908 AATACCAATGAATATATGATGCTGACATTCAGATTAACCTGAGATTAATCCACAAGAAATGT 1967
QY 1681 GTTTTGGCTCAGATCCTCTTCTGACAGAGCAAGACATGAGGGGTATGAGGCCACCATGG 1740
    |||
Db 1968 GTTTTGGCTCAGATCCTCTTCTGACAGAGCAAGACATGAGGGGTATGAGGCCACCATGG 2027
QY 1741 GGTTCACAGATGCTACATTAAGAAGCATTAATCTCCGTTGGTCTCAGAGTTAAA 1800
    |||
Db 2028 GGTTCACAGATGCTACATTAAGAAGCATTAATCTCCGTTGGTCTCAGAGTTAAA 2087
QY 1801 CCAATCAGACAGAAAAACCAAAAAGGCTGTGTGAGATACCTGATTCAGAGAGAGTGTGT 1860
    |||
Db 2088 CCAATCAGACAGAAAAACCAAAAAGGCTGTGTGAGATACCTGATTCAGAGAGAGTGTGT 2147
QY 1861 GTGAGACTGTGTAAGAGAGATGATCATCTCAAGATATGTAAGAAAGAGTTCTCAGATATCT 1920
    |||
Db 2148 GTGAGACTGTGTAAGAGAGATGATCATCTCAAGATATGTAAGAAAGAGTTCTCAGATATCT 2207
QY 1921 AGTGAATGAATATACGATCACTATTTATCCAAAATGTGTGATGAGGTTTCTCTTGCT 1980
    |||
Db 2208 AGTGAATGAATATACGATCACTATTTATCCAAAATGTGTGATGAGGTTTCTCTTGCT 2267
QY 1981 GATGAGCAACCTCAGACTTCTGACACATCAGTAGAGTACAGTTTGAACAATTACAGAA 2040
    |||
Db 2268 GATGAGCAACCTCAGACTTCTGACACATCAGTAGAGTACAGTTTGAACAATTACAGAA 2327
QY 2041 AAATCTGCGGAAAATATCAATATGCTTCAGGTTGTATCAGCTTGTAAATCT 2100
    |||
Db 2328 AAATCTGCGGAAAATATCAATATGCTTCAGGTTGTATCAGCTTGTAAATCT 2387
QY 2101 CCCAAAATCACTATTTTACAGATATGCTAAATGCAATTTGATGAGAAATTCCTCGGT 2160
    |||
Db 2388 CCCAAAATCACTATTTTACAGATATGCTAAATGCAATTTGATGAGAAATTCCTCGGT 2447
QY 2161 GCTGATTTTGAAGTTTGGTTTATGATGGGTAAATAATACAAAACAGAAATTTCAAT 2220
    |||
Db 2448 GCTGATTTTGAAGTTTGGTTTATGATGGGTAAATAATACAAAACAGAAATTTCAAT 2507
QY 2221 CAGGTGATTTGAAAAGACAGGGAAGTCTTACACTTTAAAAGTGAAGTGAATATATGC 2280
    |||
Db 2508 CAGGTGATTTGAAAAGACAGGGAAGTCTTACACTTTAAAAGTGAAGTGAATATATGC 2567
QY 2281 TTGAAAAGAGAGATAAAATATGATATGACCATGCTAATGAGGGTCAATCGATTTGTTTA 2340
    |||
Db 2568 TTGAAAAGAGAGATAAAATATGATATGACCATGCTAATGAGGGTCAATCGATTTGTTTA 2627
QY 2341 GCACCTGGAATCCATAATTTTCAGAGAGAGAAAGAAAACCTAGAGTGTCTCTTTTCCCA 2400
    |||
Db 2628 GCACCTGGAATCCATAATTTTCAGAGAGAGAAAGAAAACCTAGAGTGTCTCTTTTCCCA 2687
QY 2401 AATAATCATAGGAAGAAAACCTGTATGATCTATGTTCACTAGAGGCTTATACCTCTCTCT 2460
    |||
Db 2688 AATAATCATAGGAAGAAAACCTGTATGATCTATGTTCACTAGAGGCTTATACCTCTCTCT 2747
```

```
QY 2461 TCTGTGATTTCAATTTATCCCAACGAGATGAGCATCTTTCAACAGAAATGCTATGCAT 2520
    |||
Db 2748 TCTGTGATTTCAATTTATCCCAACGAGATGAGCATCTTTCAACAGAAATGCTATGCAT 2807
QY 2521 AGTGTGCTTTCTCCCAACAGGCAACCAATCCCTTATCCCTCTGATGAGTTAAAGAAAGGA 2580
    |||
Db 2808 AGTGTGCTTTCTCCCAACAGGCAACCAATCCCTTATCCCTCTGATGAGTTAAAGAAAGGA 2867
QY 2581 CTGTGGTCTTACACTACAGCTTCTGGAACAGACATCTCTTAATATGCTAAAAAGATGT 2640
    |||
Db 2868 CTGTGGTCTTACACTACAGCTTCTGGAACAGACATCTCTTAATATGCTAAAAAGATGT 2927
QY 2641 CTTCCTTAATCAGCACTAATCTTTTGAATCTTTTGGAAAAATGTTGGTGGCTTACA 2700
    |||
Db 2928 CTTCCTTAATCAGCACTAATCTTTTGAATCTTTTGGAAAAATGTTGGTGGCTTACA 2987
QY 2701 CAGTTAACTAGTGAAGCTGTGTGAGTTCAGTTAATGATGAGTCCAGATTGGTGTGAG 2760
    |||
Db 2988 CAGTTAACTAGTGAAGCTGTGTGAGTTCAGTTAATGATGAGTCCAGTTGGTGTGAG 3047
QY 2761 GCAGAGAGTCTTCTTATCAGTTATACCTCACCAATGCTCAACACTAGGTATGAGAA 2820
    |||
Db 3048 GCAGAGAGTCTTCTTATCAGTTATACCTCACCAATGCTCAACACTAGGTATGAGAA 3107
QY 2821 AATGAAAATTTACAGCACTACATCAACAGAAATTAAGATGTCTGTCTTCATCTTTTG 2880
    |||
Db 3108 AATGAAAATTTACAGCACTACATCAACAGAAATTAAGATGTCTGTCTTCATCTTTTG 3167
QY 2881 ATGTTTCTAATCCGACTCTTAATTTTCATTGA 2913
    |||
Db 3168 ATGTTTCTAATCCGACTCTTAATTTTCATTGA 3200

RESULT 5
AAKS2950
ID AAKS2950 standard; cDNA; 3841 BP.
XX
AC AAKS2950;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2479.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
    vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
    tissue growth factor; immunomodulatory; cancer; leukaemia;
    nervous system disorder; arthritis; inflammation; se.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM79817.
```

XX Nucleic acid encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.

XX Claim 1, Page 4757-4758; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW7823-AAW80302) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibit activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication

XX Sequence 3841 BP; 1253 A; 704 C; 764 G; 1120 T; 0 U; 0 Other;

Query Match 97.0%; Score 2825.8; DB 4; Length 3841;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 2901; Conservative 0; Mismatches 12; Indels 8; Gaps 6;

QY 1 ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTTGTAA 60  
DB 228 ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTTGTAA 287  
QY 61 GGATCATTTGCTGCTGCTCTACAGAGCTGAGTCCATTCACAGCTGTTGGAAAGTTGCAATC 120  
DB 288 GGATCATTTGCTGCTGCTCTACAGAGCTGAGTCCATTCACAGCTGTTGGAAAGTTGCAATC 347  
QY 121 AAAATGATGATTAAGAAAGCATGTA CAAGCAGAAAGTGTACAGAGAGTCCAAATGAG 180  
DB 348 AAAATGATGATTAAGAAAGCATGTA CAAGCAGAAAGTGTACAGAGAGTCCAAATGAG 407  
QY 181 GTGAAATACATTTGCCAATTTGAAACATCTTCTATCTTGGAGTTTATATCTATTTTGA 240  
DB 408 GTGAAATACATTTGCCAATTTGAAACATCTTCTATCTTGGAGTTTATATCTATTTTGA 467  
QY 241 GATGACATTTATGATGATGATGATTAAGAAATGTCATTAATGAGAAATGAAACAGTAT 300  
DB 468 GATGACATTTATGATGATGATGATTAAGAAATGTCATTAATGAGAAATGAAACAGTAT 527  
QY 301 CTAAAGATAGAGTAAACCTTCTCAGAAATGAAAGTCTGACATTCATGACACAGATC 360  
DB 528 CTAAAGATAGAGTAAACCTTCTCAGAAATGAAAGTCTGACATTCATGACACAGATC 587  
QY 361 ATTCACAGGATGTTGATCTTCTATCTGATATCTACACGGAGCTCAGACATTTCT 420  
DB 588 ATTCACAGGATGTTGATCTTCTATCTGATATCTACACGGAGCTCAGACATTTCT 647  
QY 421 AACCTCTCTAGCTGCTGATATGAAACATCAAGATTGCTGATTTGGGCTGGCACTCA 480  
DB 648 AACCTCTCTAGCTGCTGATATGAAACATCAAGATTGCTGATTTGGGCTGGCACTCA 707  
QY 481 CTGAAATGCAATGAAAGCACTATATGAGAACTCTTAATCTTCAATTTGACCA 540  
DB 708 CTGAAATGCAATGAAAGCACTATATGAGAACTCTTAATCTTCAATTTGACCA 767  
QY 541 GAAATGCACTGCAAGTGCATGAGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 600  
DB 768 GAAATGCACTGCAAGTGCATGAGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 827  
QY 601 TTTTATACATTAATGCTGAGGAGACCACTTGACATGACAGATCAAGAACATTA 660  
DB 828 TTTTATACATTAATGCTGAGGAGACCACTTGACATGACAGATCAAGAACATTA 887  
QY 661 AATTAAGATGATGTCAGATTAATGAATGCACTTTTGTCAATAGAGGCAAGAC 720  
DB 888 AATTAAGATGATGTCAGATTAATGAATGCACTTTTGTCAATAGAGGCAAGAC 947

QY 721 CTATATACCAAGTACTTGTAGAAATCCAGAGAGATGTTTAAGTCTGTCTCAATTTG 780  
DB 948 CTATATACCAAGTACTTGTAGAAATCCAGAGAGATGTTTAAGTCTGTCTCAATTTG 1007  
QY 781 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840  
DB 1008 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 1067  
QY 841 GACTCAATTTGATGAGGAGATGCAATTTTCTACTGCAATTCAGCTTCTTCCAGTAC 900  
DB 1068 GACTCAATTTGATGAGGAGATGCAATTTTCTACTGCAATTCAGCTTCTTCCAGTAC 1127  
QY 901 AGTATAGTATGATTTATTTAGCAAAAGAGACTTTGATTTGTCAGGCACTCCCAAT 960  
DB 1128 AGTATAGTATGATTTATTTATTAACAAAAGAGACTTTGATTTGTCAGGCACTCCCAAT 1187  
QY 961 AAAATGACTGATTTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAAGAGATGA 1020  
DB 1188 AAAATGACTGATTTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAAGAGATGA 1247  
QY 1021 AACAGTTTATATCTCAAGTGGGAAATCAAGAAACAGATTAATGAGAAAGGAGATTA 1080  
DB 1248 AACAGTTTATATCTCAAGTGGGAAATCAAGAAACAGATTAATGAGAAAGGAGATTA 1307  
QY 1081 ATTCAAGATGAGAAAGGACCACTCTGATACCTTCTGATGAGCTTAATCTCTGAT 1140  
DB 1308 ATTCAAGATGAGAAAGGACCACTCTGATACCTTCTGATGAGCTTAATCTCTGAT 1367  
QY 1141 AGATCTGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGAGATGTCAC 1200  
DB 1368 AGATCTGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGAGATGTCAC 1427  
QY 1201 TCAGCAAAATGCTTCAAGTCTCAAAAGATCAGAGAGAGTGAATTAAGAGAGTAC 1260  
DB 1428 TCAGCAAAATGCTTCAAGTCTCAAAAGATCAGAGAGAGTGAATTAAGAGAGTAC 1487  
QY 1261 TCAGCAAGACAAAGATGCAAACTTTTAACTTTTAAAGAAAGACATCCAGTACT 1320  
DB 1488 TCAGCAAGACAAAGATGCAAACTTTTAACTTTTAAAGAAAGACATCCAGTACT 1547  
QY 1321 TCTGATCTTTTGAAGAGCTGATTAACATCAAGACCTTCCATCATCTTTGTCCAGGA 1380  
DB 1548 TCTGATCTTTTGAAGAGCTGATTAACATCAAGACCTTCCATCATCTTTGTCCAGGA 1607  
QY 1381 AAACTCTTTTCAATTTGCAAGCCGACACTCAGACTGAAACCGTACACAGTGTCT 1440  
DB 1608 AAACTCTTTTCAATTTGCAAGCCGACACTCAGACTGAAACCGTACACAGTGTCT 1667  
QY 1441 GGAATTCGCAATTAAGTCTCAATTAAGAAAGAACTGATGATTAAGACAGATCCAGCCCA 1500  
DB 1668 GGAATTCGCAATTAAGTCTCAATTAAGAAAGAACTGATGATTAAGACAGATCCAGCCCA 1727  
QY 1501 AACCGGACCTTCAGGGCCATCAGATTTGCAAGAGGACACATCAAAAATGCTGACT 1560  
DB 1728 AACCGGACCTTCAGGGCCATCAGATTTGCAAGAGGACACATCAAAAATGCTGACT 1787  
QY 1561 GATACAAAAGTCAAAAAGAACTGATGCTTGTATATGACATTTCTGTAAACAGCAA 1620  
DB 1788 GATACAAAAGTCAAAAAGAACTGATGCTTGTATATGACATTTCTGTAAACAGCAA 1847  
QY 1621 AATACCAAGAAATTAATGATGCACTTCAACATTAACCTGATTAATCAACAAAGATGT 1680  
DB 1848 AATACCAAGAAATTAATGATGCACTTCAACATTAACCTGATTAATCAACAAAGATGT 1907  
QY 1681 GTTTTGGCTCAGATCTCTTCTGAAAGAGCAAGACTAGGGGTATGAGGCCACATAG 1740  
DB 1908 GTTTTGGCTCAGATCTCTTCTGAAAGAGCAAGACTAGGGGTATGAGGCCACATAG 1967  
QY 1741 GGTATACGAATGATCAATTAAGAGCATTAATCTCGTGTGTTCTCAACAGTTTAA 1800  
DB 1968 GGTATACGAATGATCAATTAAGAGCATTAATCTCGTGTGTTCTCAACAGTTTAA 2027

```

QY 1801 CCAATCAGACAGAAAAACCAAAAAAGCGTGTGACATCTTGTATTGAGAGAGGTGTG 1860
DB 2028 CCAATCAGACAGAAAAACCAAAAAAGCGTGTGACATCTTGTATTGAGAGAGGTGTG 2087
QY 1861 GTGAGCTTGTAAAGATATGATCTCAAGAAATGTGAAAGAAATTCTTCAGATATCT 1920
DB 2088 GTGAGCTTGTAAAGATATGATCTCAAGAAATGTGAAAGAAATTCTTCAGATATCT 2147
QY 1921 AGTGAATGAAATACGATCAATATTTATATCAAAATGTGGTAAAGGTTTCTCTTGC 1979
DB 2148 AGTGAATGAAATACGATCAATATTTATATCAAAATGTGGTAAAGGTTTCTCTTGC 2207
QY 1980 T-GAATGACCAACCTCAGCTTAC-TGACAAATCAGTAAAG--TACAGCTTGAACAATTTAC 2035
DB 2208 TGAATGACCAACCTCAGCTTGAACAATCAGTAAAGGTTTGAACAATTTAC 2267
QY 2036 CAGAAAAATATCTGCGAATAATATCAATATGCTCAGGTTTGTACAGCTTGAAGATCTA 2095
DB 2268 CAGAAAAATATCTGCGAATAATATCAATATGCTCAGGTTTGTACAGCTTGAAGATCTA 2327
QY 2096 AATCTCCCAAAATCACTTATTTTACAAGATATGCTAAATGATTTGATGAGAAATTTCTC 2155
DB 2328 AATCTCCCAAAATCACTTATTTTACAAGATATGCTAAATGATTTGATGAGAAATTTCTC 2387
QY 2156 CTGGTGTGATTTTGAAGTTTGTGTTTATGATGGGTAAATACCAAAACGAAAGATT 2215
DB 2388 CTGGTGTGATTTTGAAGTTTGTGTTTATGATGGGTAAATACCAAAACGAAAGATT 2447
QY 2216 TCATTCAGGTGATTTGAAAGACAGGAGTCTTACACTTTAAATGAAAGTAAAGTTA 2275
DB 2448 TCATTCAGGTGATTTGAAAGACAGGAGTCTTACACTTTAAATGAAAGTAAAGTTA 2507
QY 2276 ATAGCTTGAAGAGAGATTAATAATGTATATGACATGCTAAATGAGGTCATCGTATTT 2335
DB 2508 ATAGCTTGAAGAGAGATTAATAATGTATATGACATGCTAAATGAGGTCATCGTATTT 2567
QY 2336 GTTTAGACATGGAATCCATATTTTCAAGAAAGAAAGAACTAGAGATGCTCCCTTTT 2395
DB 2568 GTTTAGACATGGAATCCATATTTTCAAGAAAGAAAGAACTAGAGATGCTCCCTTTT 2627
QY 2396 TCCCAATATCAATGAAAGAAACCTGTGTATGATCTATGCTACCTTAAGGCTTATACCTC 2455
DB 2628 TCCCAATATCAATGAAAGAAACCTGTGTATGATCTATGCTACCTTAAGGCTTATACCTC 2687
QY 2456 CTCCTTCTGTGATTAATAATACCAAGAGATAGACATCTTTCACAGAAATGCTCA 2515
DB 2688 CTCCTTCTGTGATTAATAATACCAAGAGATAGACATCTTTCACAGAAATGCTCA 2747
QY 2516 TGCATGTGCTGCTTCCACACAGGACCAATCCTTAATCCCTATGTTCAAAATG 2575
DB 2748 TGCATGTGCTGCTTCCACACAGGACCAATCCTTAATCCCTATGTTCAAAATG 2807
QY 2576 AAGGACTTGTCTTCAATACAGCTTTCGAAACAGACATCTTCTAATAGTCTAAAG 2635
DB 2808 AAGGACTTGTCTTCAATACAGCTTTCGAAACAGACATCTTCTAATAGTCTAAAG 2867
QY 2636 ATGTCTTCTTAATCAGACAACTTTGAAATGTTTGTGAAATGTTGTTGGG 2695
DB 2868 ATGTCTTCTTAATCAGACAACTTTGAAATGTTTGTGAAATGTTGTTGGG 2927
QY 2696 CTACAC-AGTTAACTAGTGAAGCTGTGGGTTCAATTAATGATGGTCCAGTTGGT 2754
DB 2928 CTACACAGTTAACTAGTGAAGCTGTGGGTTCAATTAATGATGGTCCAGTTGGT 2997
QY 2755 GTGACGAGAGAGTCTTCTATCAGTATATCTCACCAAAATGTTGTTGTTG 2812
DB 2998 GTGACGAGAGAGTCTTCTATCAGTATATCTCACCAAAATGTTGTTGTTG 3047
QY 2813 ATGAGAAAAATGAAAAATTTACCAAGCTATCAATCAAAAGAAATTTACAGTGTCTCTTCCA 2872
DB 3048 ATGAGAAAAATGAAAAATTTACCAAGCTATCAATCAAAAGAAATTTACAGTGTCTCTTCCA 3107
QY 2873 TCCTTTGATGTCTTCTAATCCGACTCTCTAATTTTCAATGA 2913

```

```

DB 3108 TCCTTTGATGTCTTCTAATCCGACTCTAATTTTCAATGA 3148

RESULT 6
AAI60186
ID AAI60186 standard; cDNA, 3845 BP.
XX
XX AAI60186;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4175.
XX
XX Human, nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX chemokine; lateral sclerosis; Shy-Drager Syndrome; chemokine;
XX leukemia; ss.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-0048725.
XX 25-APR-2000; 2000US-0052317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX
XX (HYSB-) HYSB INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41030.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4175; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytoskeletal activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 3845 BP; 1253 A; 712 C; 764 G; 1116 T; 0 U; 0 Other;
XX
XX Query Match 97.0%; Score 2825.8; DB 4; Length 3845;
XX Best Local Similarity 99.3%; Pred. No. 0;

```



QY 2156 CTGGTCTGATTTTGAAGTTTGGTTTATGATGGGGTAAAAATACAGAAACAGAGATT 2215  
 Db 2388 CTGGTCTGATTTTGAAGTTTGGTTTATGATGGGGTAAAAATACAGAAACAGAGATT 2447  
 QY 2216 TCATTGAGTGAATTTAAAAAGCAGGAAAGTTTACCTTTAAAAAGTAAAGTAACTTA 2275  
 Db 2448 TCATTGAGTGAATTTAAAAAGCAGGAAAGTTTACCTTTAAAAAGTAAAGTAACTTA 2507  
 QY 2276 ATAGCTTGAAGAGAGATTTAAAAATGTATATGACATGCTAAATGAGGGTCACTGTTT 2335  
 Db 2508 ATAGCTTGAAGAGAGATTTAAAAATGTATATGACATGCTAAATGAGGGTCACTGTTT 2567  
 QY 2336 GTTTAGACATGGAATTCCTAATTTTCAAGAGAGAAAAGTAACTGAGAGTCTCCCTTT 2395  
 Db 2568 GTTTAGACATGGAATTCCTAATTTTCAAGAGAGAAAAGTAACTGAGAGTCTCCCTTT 2627  
 QY 2396 TCCCAATTAATCATAGAAAGAAACCTGTGTATCTAGTTTACCTAAGGCTTTATCACCTC 2455  
 Db 2628 TCCCAATTAATCATAGAAAGAAACCTGTGTATCTAGTTTACCTAAGGCTTTATCACCTC 2687  
 QY 2456 CTCCTTCTGTGATTTCAAAATTAACCAACGAGATAGAGATCTTTCAACAGAAATGCTCA 2515  
 Db 2688 CTCCTTCTGTGATTTCAAAATTAACCAACGAGATAGAGATCTTTCAACAGAAATGCTCA 2747  
 QY 2516 TGGATGTGCTGCTCTCCCAACAGGACCAATCTTAACTCCCTTATGTGTTACAAATG 2575  
 Db 2748 TGGATGTGCTGCTCTCCCAACAGGACCAATCTTAACTCCCTTATGTGTTACAAATG 2807  
 QY 2576 AAGGACTTGTCTTAACTAAGTCTTCTGGAACAGACATCTCTTAAATAGCTTAAAG 2635  
 Db 2808 AAGGACTTGTCTTAACTAAGTCTTCTGGAACAGACATCTCTTAAATAGCTTAAAG 2867  
 QY 2636 ATTGTCTTCTTAAATCAGCAAACTTTTGAATCTGTGTTTGTGAAAATGTGTGTTGGG 2695  
 Db 2868 ATTGTCTTCTTAAATCAGCAAACTTTTGAATCTGTGTTTGTGAAAATGTGTGTTGGG 2927  
 QY 2696 CTACAC-AGTTAATCTAGGAGCTGTGTGAGGTTTCACTTTAAATAGTGGTCCAGTTGGT 2754  
 Db 2928 CTACACAGTTAATCTAGGAGCTGTGTGAGGTTTCACTTTAAATAGTGGTCCAGTTGGT 2987  
 QY 2755 GTGACAGGAGAGAGTCTTCTATCACTTAACTTCACTCAAAATGTGCAAA--CAACTAAGT 2812  
 Db 2988 GTGACAGGAGAGAGTCTTCTATCACTTAACTTCACTCAAAATGTGCAAACTAAGT 3047  
 QY 2813 ATGAGAGAAATGAAATTAACGAGCTTACATCAAAAGAAATTAAGTGTCTGCTTCA 2872  
 Db 3048 ATGAGAGAAATGAAATTAACGAGCTTACATCAAAAGAAATTAAGTGTCTGCTTCA 3107  
 QY 2873 TCCCTTTGATGTTTCTAATCCGACTCTCTAATTTTCAATTTGA 2913  
 Db 3108 TCCCTTTGATGTTTCTAATCCGACTCTCTAATTTTCAATTTGA 3148

RESULT 7  
 AB19713  
 ID AB19713 standard; cDNA; 3447 BP.  
 AC AB19713;  
 XX  
 XX 07-MAR-2002 (first entry)  
 DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:765.  
 XX  
 XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.  
 XX  
 OS Mus musculus.  
 XX  
 XX W0200188188-A2.  
 XX  
 XX 22-NOV-2001.  
 PD  
 XX 18-MAY-2001; 2001WO-JP004192.  
 PF

XX  
 PR 18-MAY-2000; 2000CP-00145977.  
 XX  
 PA (UNYI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
 XX  
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
 XX  
 DR WPI; 2002-03473/04.  
 XX  
 DR P-PSDB; ABB57273.  
 XX  
 PS Claim 2; Page 1887-1893; 2690pp; English.  
 XX  
 CC The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (AB199202 to AB199912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 3447 BP; 1061 A; 690 C; 754 G; 942 T; 0 U; 0 Other;  
 Query Match 64.5%; Score 1879; DB 6; Length 3447;  
 Best Local Similarity 80.9%; Pred. No. 0;  
 Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;  
 QY 1 ATGGCGACCTGCACTGGGAGAGAAATGAGAGATTTTAAATTTGAATCTGCTTGTAA 60  
 Db 206 ATGGCGGCGTGCAGCGGAGAGATCGAGACTTTAAAGTTGAAATCTACGCGTAA 265  
 QY 61 GGATCATTTGCTGTGTCTTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 120  
 Db 266 GGATCATTTGCTGTGTCTTACAGAGCTGAGTCCATTCACACTGGTTTGAAGTTGCAATC 325  
 QY 121 AAAATGATTAATGAAGAGCATCTACAAAGCAAGAAATGTAACAGAGTCCAAATGAG 180  
 Db 326 AAAATGATTAATGAAGAGCATCTACAAAGCAAGAAATGTAACAGAGTCCAAATGAG 385  
 QY 181 GTGAAATATACATTCGCAATTTGAACATCTCTATCTGTGAGCTTTATTAATTTTGA 240  
 Db 386 GTGAAATATACATTCGCAATTTGAACATCTCTATCTGTGAGCTTTATTAATTTTGA 445  
 QY 241 GATAGCAATTAATGTATCTGTATTGAATGTGCAATATGAGAAATGAAACAGATAT 300  
 Db 446 GATAGCAATTAATGTATCTGTATTGAATGTGCAATATGAGAAATGAAACAGATAT 505  
 QY 301 CTAAGAAATGAGTGAAGCCCTTCTGAGAAATGAAAGTCTGACACTTTCACCAATC 360  
 Db 506 CTAAGAAATGAGTGAAGCCCTTCTGAGAAATGAAAGTCTGACACTTTCACCAATC 565  
 QY 361 ATCAGAGGATGTTTATCTTCAATTTCTAGTAACTACACCGGAGCTCACATTTT 420  
 Db 566 ATCAGAGGATGTTTATCTTCAATTTCTAGTAACTACACCGGAGCTCACATTTT 625  
 QY 421 AACCTCTTCTGACTGTGTAATTAATGAACATCAAGATTCGATTTTGGCTGGCACTCA 480  
 Db 626 AACCTCTTCTGACTGTGTAATTAATGAACATCAAGATTCGATTTTGGCTGGCACTCA 685  
 QY 481 CTGAAATGCAATGAAGAGCATATCATTAATGGAATCTCCCTTAATCACTTACCA 540  
 Db 686 TTGAATATGCCAATGAAGAGCATATCATTAATGGAATCTCCCTTAATTAATTTACCA 745

QY 541 GAAATTCGCACTGCAAGTGCACATGGCCCTTGAACTGATGTTTGGTCCCTGGCTGTAG 600  
DB 746 GAAATTCGCACTGCAAGTGCACATGGCACTTGAACTGATGTTTGGTCCCTGGCTGTAG 805  
QY 601 TTTTATACATTAATTATCCGGAGACCACTTCGACATGACAGCTGCAAGACATTA 660  
DB 806 TCTTATACATTAATTATCCGGAGACCACTTCGACATGACAGCTGCAAGACATTA 865  
QY 661 AATTAAGTATGATGGCAGATTAATGAATGCCATCTTTTGTCAATGAGGCCAAGAC 720  
DB 866 AACAAAGTGTCTCCGCAATTAATGAATGCCAGCTTTTGTCAATGAGGCCAAGAC 925  
QY 721 CTTATTCACAGCTTCTTGTGTAATCCAGCAGATGCTTAACTGCTTCAATGATG 780  
DB 926 CTTATTCACAGCTTCTTGTGTAATCCAGCAGATGCTTAACTGCTTCAATGATG 985  
QY 781 GACCATCTTTTATGCTCCGAAATCTTCAACAAAAGTAAGATTAGAACTGTGAA 840  
DB 986 GACCATCTTTTATGCTCCGAAATCTTCAACAAAAGTAAGATGAGCTGTAGAG 1045  
QY 841 GACTCAATGATGATGGGCAATGCCCAATTTCTACCTGCAATTAAGCTTCTTCACTACC 900  
DB 1046 GACTCAATGATGATGGGCAATGCCCAATTTCTACCTGCAATTAAGCTTCTTCACTACC 1105  
QY 901 AGTATTAAGTATGATTTATTTGACAAAAGACCTTTGATGCTGACGCACTCCCAAT 960  
DB 1106 AGTATTAAGTATGATTTATTTGACAAAAGACCTTTGATGCTGACGCACTCCCAAT 1162  
QY 961 AAAATGACTGATATTTTCAAGAATAAAGTTCACATGATTTTCTTCTTCAAGAGATGA 1020  
DB 1163 AAAATGACTGATATTTTCAAGAATAAAGTTCACATGATTTTCTTCTTCAAGAGATGA 1219  
QY 1021 AACAGTTTATATCTCAGTGGGGAAT-----CAAGAAACCAATATAGTGAAGGGA 1074  
DB 1220 AATAATTTTATATCTCAGTGGGGAATCCAGAAACAAGATATAGTGAAGGGA 1279  
QY 1075 AAGATTAATCAAGATGCAAGAAAGGCAATTCCTGACATCCTTGTAGCTTATTC 1134  
DB 1280 AAGATTAATCAAGATGCAAGAAAGGCAATTCCTGACATCCTTGTAGCTTATTC 1339  
QY 1135 TCTGATAGATCTGGCACTCTAATAGTCACTGCTCAAGCAAAAACATATCAATGGAACA 1194  
DB 1340 TCTGATAGATCTGGCACTCTAATAGTCACTGCTCAAGCAAAAACATATCAATGGAACA 1396  
QY 1195 TGTCACTGACAGAAATGCTTCAAGTGCMAAAGATCAGAGAGAGGTGAAGAG 1254  
DB 1397 TGTCACTGACAGAAATGCTTCAAGGCTTCAAGAGATC----- 1434  
QY 1255 AGTACTCACCAAGACAAATGCCMACTTTTAACTTCTTAAAGAAAAGACATCC 1314  
DB 1435 ----- 1434  
QY 1315 AGTATTTCTGATCTTTGAAAAGCTGATTAACATCAAGCACTCTCAATCATCTTGT 1374  
DB 1435 -----ACTGATGAAAATCAACACAGTTCCATCATCTTGT 1471  
QY 1375 CCAGAAAAATCTCTTTTCATTTGCAAGCCGACACTCACTGAAAACCGTAAACAG 1434  
DB 1472 CTAGAAAAATCTCTTTTCATTTGCAAGCCGACACTCACTGAAAATGTTAGACAGAG 1531  
QY 1435 TGTGTTGGGAATCTGAAAATGATGCTATTTAGGAAAATATATAGCAACCAACGTT 1591  
DB 1532 TGTGTTGGGAATCTGAAAATGATGCTATTTAGGAAAATATATAGCAACCAACGTT 1591  
QY 1495 AGCCCAAAACCGGGACTTCAGAGGCACTGATTTGCAAGAGACATCAAAAATGCGC 1554  
DB 1592 AGCCCAAAACCGGGAATTTCCAGGCACTTCCAGATTTGC--AGACACGTTTACAAAACGCT 1648  
QY 1555 TGGACTGATTAACAAGATCAAAAAGAACTGTGATGCTTGTGATTAATGCAATTTCTGTA 1614  
DB 1649 TGGACTGATTAACAAGATCAAAAAGAACTGTGATGCTTGTGATTAATGCAATTTCTGTA 1708

QY 1615 CAGCAAAATACATGAAAATATATGACTGCACTTCAAGTAAACCTGATTAATCCAA 1674  
DB 1709 CAGCTGAGTGCATGAAATACATGATGCACTTCAACATTAACCTGAGGCTCATGCAAG 1768  
QY 1675 GAATGTTTTTGGCTCAGATCTCTTTCTGAAACAGAGCAAGACTTGGGATATGAGCCA 1734  
DB 1769 GA-----GCCGGGCTCATCTCATCTTGAACAAAGCAAAATGAAATGAGTGTG 1822  
QY 1735 CCATGGGCTTATCAGAAATGCTATTTAAGAAACATTAACCTCCGTTGGTGTCTACAG 1794  
DB 1823 AACTGGGCTTATCAGAAATGCTATTTAAGAAACATTAACCTCCGTTGGTGTCTACAG 1882  
QY 1795 TTTAAACCAATCAGACAGAAAAACAAAAGCTGTGTGAGCACTTGAATCAGAGAG 1854  
DB 1883 TTTAAACCAATCAGACAGAAAAACAAAAGCTGTGTGAGCACTTGAATCAGAGAG 1942  
QY 1855 GTGTGTGTGAGCTTGTAAAGAGTATGATCTCAAGAAATATGTGAAGAGTTCTTCA 1914  
DB 1943 GTGTGTGTGAGCTTGTGAAGAGTGTGCTGTGAAGATATGTGAAGAGTTCTTCA 2002  
QY 1915 AATATGATGATGAAATACGATCACTATTTATTTATCCAAATGATGATGATTTTCT 1974  
DB 2003 AATATGATGATGATGAAATACGATCACTATTTATTTATCCGAAAGATGAAAGGCTTCT 2062  
QY 1975 CTGCTGATGACCAACCTCACTTCTGACAAATCAGTATGATGACGCTTGAATTA 2034  
DB 2063 CTGCTGACAGACTTCTGCTGCTTACTGACAAATCAGTATGATGACGCTTGAATTA 2122  
QY 2035 CCAGAAAAATCTGCGGAAATATCAATATGCTTCCAGGTTGTACAGCTTGTATGATCT 2094  
DB 2123 CCAGAAAAATCTGCGGAAATATCAATATGCTTCCAGATTCATTCAGCTTGTATGATCT 2182  
QY 2095 AATATCCCAAAATCACTATTTTATCAAGATATGCTTAAATGCAATTTGATGAGAAATCT 2154  
DB 2183 AAAAACTCCCAAAATCACTATTTTATCAAGATATGCTTAAATGCAATTTGATGAGAAATCT 2242  
QY 2155 CCGTGTGCTGATTTTGAAGGTTGTTTATGATGAGGTTTAAATATACAAACAGAAAGAT 2214  
DB 2243 CCGTGTGCTGATTTTGAAGGTTGTTTATGATGAGGTTTAAATATACAAACAGAAAGAT 2302  
QY 2215 TTTCAATCAGTATGTAAGAAAGACAGGAGTCTTACATTTTAAAGTGAAGTGAAT 2274  
DB 2303 TTTCAATCAGTATGTAAGAAAGACAGGAGTCTTATATTTTAAAGTGAAGTGAAT 2362  
QY 2275 AATATGTAAGAAAGAGATTAATAATATATGACATGCTTAAAGAGGCTCATGAT 2334  
DB 2363 AATATGTAAGAAAGAGATTAATAATATATGACATGCTTAAAGAGGCTCATGAT 2422  
QY 2335 TGTATGACAGTGAATCCATTAATTTCAAGAGAGAAAGAAACATGAGAGTCTCCCTT 2394  
DB 2423 TGTATGACAGTGAATCTGTAATCTTGAAGAGAAAGAAAGAGAGGCTTCTTCAATTC 2482  
QY 2395 TTTCCAAATATCATGAAAGAAACCTGTGATGATGATTTCACTTAAAGCTTATCACT 2454  
DB 2483 TTTCCAAATATCATGAAAGAAACCTGTGATGATGATTTCACTTAAAGCTTATCACT 2542  
QY 2455 CCTCTCTTGTGATGATTAATTAACCAAGAGATGAGCATTTTCAACGAATGCTC 2514  
DB 2543 CCTC---CTGTGAGCCCAAGCTGTGTAAAGAGAGAGGCTGTCAAGAGAGAGT 2599  
QY 2515 ATGATATGATGCTTCTTCCAAACAGAGCAACATCTTAACTCCCTTATGTTATCAAT 2574  
DB 2600 ATGATATGATGCTTCTTCCCAACAGAGCTCCCAAGAGCTCAAGTCTTCACTGAGAG 2659  
QY 2575 GAAGACTTGTCTTAACTCACTCAAGCTTGTGAACAGACATCTTCTTAAATGCTTAA 2634  
DB 2660 GAAGACTTGTGTCAACAGAGCTGTGCAAGAGAGGCTGTCTTCA----- 2708  
QY 2635 GATGCTCTTCTTAAATCAGACAACTTTTGAATCTGTTTGTGAAAATATGTTGTTGG 2694  
DB 2709 ---GCTTCTTAAATCTGCAAGCTTTTGAATCTGTTTGTGAAAATATGTTGTTGG 2764  
QY 2695 GCTACACAGTAACTATGATGAGCTGTGTGAGTTCAATGATGAGTCCAGTTGTT 2754

```

Db      2765 GCTACACAGCTAACTACGAGAGCTGTGGTTCACTTATATGAGTGGTCACTGGTGT 2824
QY      2755 GTGACAGCAGAGAGCTCTTCTATATGATTAACTCAACCAATGGTCAACACTAGGTAT 2814
Db      2825 GTCCAGGACAGAGATATCTTCATCATGATTACATCAATCAGATGGTGCAGACACTAGGTAT 2884
QY      2815 GAGAAATATGAAATATACCACTACATCAATCAAGAAATTAAGTGTCTGTCTTCATC 2874
Db      2885 GAGAAATATGAAATATACCTGAATACATCAATCAAGAAATTAAGTGTCTGTCTTCATC 2944
QY      2875 CTTTGTATGTTTCTTAATCCGACTCTTAATTTTCATTTGA 2913
Db      2945 CTTCTGATGTTTCTTAATCCACTCTTAATTTTCATTTGA 2983

```

```

RESULT 8
AAT08711
ID      AAT08711 standard; cDNA; 3447 BP.
AC      AAT08711;
XX      25-MAY-1996 (first entry)
XX      Sak-a serine-threonine kinase cDNA.
DE      Sak-a; serine-threonine kinase; agonist; antagonist;
KW      proliferative disease; cancer; tumour; antisense; transgenic animal;
XX      therapy; ss.
XX      Mus musculus.
OS      Mus musculus.
FH      Key Location/Qualifiers
FT      CDS 206..2983
FT      /tag= a
FT      /note= "T bases may also be U"
FT      misc_difference 2767
FT      /tag= b
FT      /note= "base n at position 2767 is unclear in the
FT      specification; encoded amino acid is Ala"
XX      CA2150789-A.
XX      03-DEC-1995.
XX      01-JUN-1995; 95CA-02150789.
XX      02-JUN-1994; 94US-00252995.
XX      (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX      Dennis JW, Heffernan M, Fode C;
XX      WPI; 1996-129817/14.
XX      P-PSDB; AAR92177.
XX      Nucleic acid encoding Sak serine-threonine kinase - useful for
XX      identifying modulators potentially useful in treatment or prevention of
XX      proliferative disease.
XX      Claim 5; Page 50-55; 73pp; English.
XX      A cDNA sequence (AAT08711) encodes sak-a (AAR92177), a novel
XX      serine/threonine kinase associated with mitotic and meiotic cell
XX      division. It was isolated from a murine lymphoid cell (D3JW25) cDNA
XX      library established in CHO cells by selection of wheategerm agglutinin-
XX      resistant clones. Another cDNA clone (AAT08712) encoded the sak-b isoform
XX      (AAR92114). Sak-a and sak-b are probably alternatively spliced forms of
XX      the gene. The sak gene can be used for prodn. of recombinant SAK, as a
XX      probe in the diagnosis of proliferative disorders or, in antisense form,
XX      may be used to treat such disorders
XX      Sequence 3447 BP; 1061 A; 690 C; 754 G; 941 T; 0 U; 1 Other;

```

```

Query Match      64.5%; Score 1878; DB 2; Length 3447;
Best Local Similarity 80.9%; Pred. No. 0;
Matches 2361; Conservative 0; Mismatches 411; Indels 147; Gaps 9;

QY      1 ATGGGCACTGATATGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTGTA 60
Db      206 ATGGGCGGCTGATGGGAGAGATCGAGACTTTAAGTTGGAATCTCTCGTAA 265
QY      61 GGATCAATTTGCTGTCTACAGAGCTAGTCACTTCACTGCTTGGAAAGTTGCATC 120
Db      266 GGATCAATTTGCTGTCTACAGAGCTAGTCACTTCACTGCTTGGAAAGTTGCATC 325
QY      121 AAAATGATGATAGAAAGCCATGTACAAAGCAGAAATGTAAGAGAGTCCAAATAG 180
Db      326 AAAATGATGATAGAAAGCCATGTACAAAGTGAATGTATAGAGAGTCCAAATAG 385
QY      181 GTGAAATATACATTCGCAATTTGAAATCCCTCATCTTGGAGCTTTATATCTATTGAA 240
Db      386 GTGAAATATACATTCGCAATTTGAAATCCCTCATCTTGGAGCTTTATATCTATTGAA 445
QY      241 GATAGCAATTAATGTATCTGTATTTGAAATGTGCCATAATGAGAAATGAAGATAT 300
Db      446 GATAGCAATTAATGTATCTGTATTTGAAATGTGCCATAATGAGAAATGAAGATAT 505
QY      301 CTAAAGATAGAGGAAACCTTCTCAGAAATGAGAGTGCACCTTCATGACCAAGATC 360
Db      506 CTGAGGAAAGAAATGAGGCTTTCTCAGAAAGGAGAGCTTACATGACCAAGAT 565
QY      361 ATCAGAGGATGTGTATCTTCAATTCATGATGATATGATACACCGGAGCTCAGACTTCT 420
Db      566 ATCAGAGGATGTGTATCTTCAATTCATGATGATATGATACACCGGAGCTCAGACTTCT 625
QY      421 AACCTCTACGATCTGTAATATGAATCAATCAAGATTCTGATTTTGGCTGGCAATCA 480
Db      626 AACCTCTACGATCTGTAATATGAATCAATCAAGATTCTGATTTTGGCAATCA 685
QY      481 CTGAAATGCGCAATGAAAGCACTATACATTTATGAGAACTCCTATCTATCTTCA 540
Db      686 TTGAATATGCGCAATGAAAGCACTATACATTTATGAGAACTCCTATCTATCTTCA 745
QY      541 GAAATGCCACTGAGATGACATGGCTTGAATCTGATGTTTGGCTGGCTGATG 600
Db      746 GAAATGCCACTGAGATGACATGGCTTGAATCTGATGTTTGGCTGGCTGATG 805
QY      601 TTTATACATTTATCTGAGGAGACCACTTGGACATCTACATGATTTGATGATGGCTGATG 860
Db      806 TTTATACATTTATCTGAGGAGACCACTTGGACATCTACATGATTTGATGATGGCTGATG 865
QY      661 AATTAAGTATGATGAGATTAATGAATGCAATCTTTTTCATTAATGAGGCAAGAG 720
Db      866 AATTAAGTATGATGAGATTAATGAATGCAATCTTTTTCATTAATGAGGCAAGAG 925
QY      721 CTATATACAGATTAATCTTGTAGAAATCCAGCATGCTTTAATGCTGTCTTCAATG 780
Db      926 CTATATACAGATTAATCTTGTAGAAATCCAGCATGCTTTAATGCTGTCTTCAATG 985
QY      781 GACCAATCTTTATGTCGCAATTTCTTCAACAAAGATTAAGATTTAGAACTGAGAA 840
Db      986 GACCAATCTTTATGTCGCAATTTCTTCAACAAAGATTAAGATTTAGAACTGAGAA 1045
QY      841 GACTCAATTTGATGAGGAGACATGCAATTTCTATGCAATTAACAGCTTCTTCAGTAC 900
Db      1046 GACTCAATTTGATGAGGAGACATGCAATTTCTATGCAATTAACAGCTTCTTCAGTAC 1105
QY      901 AGTATAGAGTATGATTTTGAACAAAGAACTTTGATGTTGTCAGCACTCCCAAT 960
Db      1106 AGTATAGAGTATGATTTTGAACAAAGAACTTTGATGTTGTCAGCACTCCCAAT 1162
QY      961 AAAATGATGATTTTCCAAAGATTAAGATTAAGTCAATGATTTTCTTCTCAGAGATGA 1020
Db      1163 AAAATGATGATTTTCCAAAGATTAAGATTAAGTCAATGATTTTCTTCTCAGAGATGA 1219

```

1021 AACAGTTTATACAGTGGGGAAT-----CAAGAAACAGTAATAGTGAAGGGA 1074  
1220 AGTAATTTTGTACTCAATGGGGAATCCAGAAACAGAAAGCTAATAGTGGGACGGGG 1279  
1075 AGAGTAATTCAGATGACAGAAAGGCAATTCGATACCTTGTAGAGCTTAATTC 1134  
1280 AAGATGATGAGATGACAGAAAGGCGCATTCGATACCTGCGAGAGCTCAATTC 1339  
1135 TCTGATAGTCTGGCACTTCTAATAGTCAAGTCCAGAGCAAAACATATCAATGAGCA 1194  
1340 TCTGATAGGACCGCTCTTA--TCAGTCTCGACAAAACATACAGAGAGCT 1396  
1195 TGTCACTCAGCAAAATGCTTCAAGTCCAAAGATCAGAGAGGTGAATGAAAG 1254  
1397 TGTCACTCAGTGAATGCTTCAAGCCTAGAGATC----- 1434  
1255 AGTACTCAACCAAGACAAATGCCAATTTTACTTCTTAAAGAAAGACATCC 1314  
1435 ----- 1434  
1315 AGTACTCTGATCTTTGAAAAGCCTGATTAACATCAAGCACTCCATCATCTTTGT 1374  
1435 -----ACTGATGAAAATCAACAGTTTCAATCATTTGT 1471  
1375 CCAGGAAAACTCTTTTCAATTTGACAGCCGACACTCAGACTGAACCGTACAAAG 1434  
1472 CTAGGAAAACTCTTTTCAATTTGACAGCCGACACTCAGATGAAAATGATAGCAG 1531  
1435 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGAAAATCACTGAATATGACAGATC 1494  
1532 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGAAAATCACTGAATATGACAGATC 1591  
1495 ACCCCAAACCGGACCTTCCAGGCGCATCCAGATTTGCAAGAGACATCAAAAATGCGC 1554  
1592 ACCCCAAAGAGATTTCCAGACATATCCAGATTTGC---AGGACAGTTACGAAAGCT 1648  
1555 TGGATGATACAAAAGTCAAAAAGACTGATGCTTCTGATTAATGCAATCTGTAA 1614  
1649 TGGATGATACAAAAGTCAAAAAGACTGATGCTTCTGATTAATGCAATCTGTAA 1708  
1615 CAGCAAAATACAGTGAATATATGACTGCACTTCAAGTAACTGAGATTAATCAACAA 1674  
1709 CAGCTGAGTCCATGAAATACATGATGACATCAACATAGCTGAGTCAATGCGCAG 1768  
1675 GATGATGTTTGGCTCAAGATCTCTTTCTGAAGACAGCAAGCTGGGATAGAGCA 1734  
1769 GA-----GCGGCGCTACATCTCTCATTCGAAACAAAGCAAGATAGAGTGG 1822  
1735 CAGTGGGTTTACAGAAATGATACATTAAGAGCAATTCCTGTTGGTGTCAAG 1794  
1823 AATGAGGTTTACAGAAATGATACATTAAGAGTATTAATCTCTGTTGCTCAAG 1882  
1795 TTAAGAACCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCACTATTGACAGAG 1854  
1883 TTAAGAACCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCACTATTGACAGAG 1942  
1855 GTGTGTGTGAGCTTGTAAAGAGTATGCACTCAAGATATGTGAAGAGTCTTCAG 1914  
1943 GTGTGTGTGAGCTTGTGAAGAGTGTGTGTGAAGATGTGAAGAGTCTTCAG 2002  
1915 AATATGATGATGAAATAGATCACTATTATTAATCAAAATGAGTGTGAGGTTTCT 1974  
2003 AATATGATGATGAAATAGATCACTATTATTAATCAAAATGAGTGTGAGGTTTCT 2062  
1975 CTGTGTGATGACACCTCACTACTGACAAATCAGATAGTGAAGCTTTGCAATTTA 2034  
2063 CTGTGTGATGACACCTCACTACTGACAAATCAGATAGTGAAGCTTTGCAATTTA 2122  
2035 CAGAAAAATATCTGGGAAAAATATCAATATGCTTCAAGTGTGAAGCTTTGAAGATCT 2094  
2123 CAGAAAAATATCTGGGAAAAATATCAATATGCTTCAAGTGTGAAGCTTTGAAGATCT 2182  
2095 AATCTCCCAAAATCACTATTATTAATCAAAATATGATTAATGATTTGATGAGAAATCT 2154

2183 AAAAATCCCAAAATCACTATTATTAACAGATATGCTAAATGATTTGATGAAAAATCT 2242  
2155 CCGTGTGATGATTTTGAAGTTTGTATTAATGATGGGTTAAATACAAACAGAAAGT 2214  
2243 CCGTGTGATGATTTTGAAGTTTGTATTAATGATGGGTTAAATACAAACAGAAAGT 2302  
2215 TTCAATCAGTATGAAAGACAGGAGTCTTCACTTTTAAAGATGAAAGTGAAGT 2274  
2303 TTCAATCAGTATGAAAGACAGGAGTCTTCACTTTTAAAGATGAAAGTGAAGT 2362  
2275 AATAGCTGAAAGAGAGATTAATAATGATATGACATGCTAATGAGGTCATCTATT 2334  
2363 ACCAGCTGAAAGAGAGATTAATAATGATATGACATGCTAATGAGGTCATCTATT 2422  
2335 TGTATGACCTGGAATCCATATTTCAAGAGAGAAAGAAACATGAGAGTCTCCCTT 2394  
2423 TGTATGACCTGGAATCTGTAATCTCTGAGAGAGAAAGAAAGAGAGGTTCTTCATT 2482  
2395 TTCCCAATATATGAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGAT 2454  
2483 TTCCCAATATATGAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGAT 2542  
2455 CCTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514  
2543 CCGTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2599  
2515 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
2600 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659  
2575 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634  
2660 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2708  
2635 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694  
2709 -----GTTCTCTTAATCTGACAGCTTTGAATCTGTTTGTGAAGATGTTGGTGG 2764  
2695 GCTACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2754  
2765 GCTACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2824  
2755 GTGACGAGAGAGTCTTCTATCAGTTATACCTGACCAATGATGATGATGATGAT 2814  
2825 GTGACGAGAGAGTCTTCTATCAGTTATACCTGACCAATGATGATGATGATGAT 2884  
2815 GAGAAATGAAATTAACAGATTAATCAATCAAGAAATTAACAGTGTCTGCTTCATC 2874  
2885 GAGAAATGAAATTAACAGATTAATCAATCAAGAAATTAACAGTGTCTGCTTCATC 2944  
2875 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2913  
2945 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2983

RESULT 9  
AAT08712  
ID AAT08712 standard; cDNA, 1599 BP.  
XX  
XX AAT08712;  
XX  
XX 25-MAY-1996 (first entry)  
XX  
XX DE Sak-b serine-threonine kinase cDNA.  
XX  
XX Sak-b; serine-threonine kinase; agonist; antagonist;  
XX KM proliferative disease; cancer; tumour; antisense; transgenic animal;  
XX therapy; ss.  
XX OS  
XX Mub musculus.  
XX  
XX FH Key Location/Qualifiers

```

FT CDS 206..1599
FT /*tag= a
FT /note= "T bases may also be U"
XX CA2150789-A.
XX
XX 03-DEC-1995.
XX
XX 01-JUN-1995; 95CA-02150789.
XX
XX 02-JUN-1994; 94US-00252995.
XX
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX Dennis JW, Heffernan M, Fode C;
XX
XX WPI; 1996-129817/14.
XX
XX P-PSDB; AAR92214.
XX
XX Nucleic acid encoding Sak serine-threonine kinase - useful for
XX identifying modulators potentially useful in treatment or prevention of
XX proliferative disease.
XX
XX Claim 6; Page 58-61; 73pp; English.
XX
XX A cDNA sequence (AAT08712) encodes sak-b (AAR92214), a novel
XX serine/threonine kinase associated with mitotic and meiotic cell
XX division. It was isolated from a murine lymphoid cell (D3JW25) cDNA
XX library established in CHOP cells by selection of wheatgerm agglutinin-
XX resistant clones. Another cDNA clone (AAT08711) encoded the sak-a isoform
XX (AAR92177). Sak-a and sak-b are probably alternatively spliced forms of
XX the gene. The sak gene can be used for prodn. of recombinant SAK, as a
XX probe in the diagnosis of proliferative disorders or, in antisense form,
XX may be used to treat such disorders
XX
XX Sequence 1599 BP; 481 A; 342 C; 378 G; 398 T; 0 U; 0 Other;
XX
Query Match 33.3%; Score 970.4; DB 2; Length 1599;
Best Local Similarity 85.0%; Pred. No. 3.5e-242;
Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;
QY 1 ATGGCAGCTGCGATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCGTGTGTAA 60
DB 206 ATGGGGCGGTGATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCGTGTGTAA 265
QY 61 GATCATTTTGTGTGTCTTACAGAGCTGATTCATCACTGTGTTGGAGTTGCAATC 120
DB 266 GATCATTTTGTGTGTCTTACAGAGCTGATTCATCACTGTGTTGGAGTTGCAATC 325
QY 121 AAAATGATGATTAAGAAAGCCATGTACAAGAGAAATGATTCAGAGAGTCCAAATGAG 180
DB 326 AAAATGATGATTAAGAAAGCCATGTACAAGAGAAATGATTCAGAGAGTCCAAATGAG 385
QY 181 GTGAAATATACATTTGCCAATGGAACATCTTATCTTGAAGCTTTATTAATCTTTTGA 240
DB 386 GTGAAATATACATTTGCCAATGGAACATCTTGAAGCTTTATTAATCTTTTGA 445
QY 241 GATAGCATTTATGTGTATCTGTATTAAGAAATGTGCCAATATGAGAAATGAACAGTAT 300
DB 446 GATAGCATTTATGTGTATCTGTATTAAGAAATGTGCCAATATGAGAAATGAACAGTAT 505
QY 301 CTAAAGATAGAGTGAACCCCTTCTCAGAAATGAAAGCTGACCTTCATGACCAAGATC 360
DB 506 CTGAAGAAACAGATGAACCTTTCTCAGAAAGGAAAGCTTGAAGCTTCATGACCAAGAT 565
QY 361 ATCAGAGGATGTGTATCTTCAATCTCATGTATACTACACCGGACCTTCACTTTCT 420
DB 566 ATCAGAGGATGTGTATCTTCAATCTCATGTATACTACACCGGACCTTCACTTTCT 625
QY 421 AACCTCCATGACTCGATTAATGAACATCAAGATGTGATTTTGGCTGGCACTCA 480
DB 626 AACATCTTACTTAAGCGGAAATGAACATTAATAATTTGCTGACTTTGGAAGTCAAGCAG 685

```

```

QY 481 CTGAAATATCCCATGAAAGCACTATACATTAATGTGAACCTCTTAACATTTTCACCA 540
DB 686 TTGAATATCCCATGAAAGCACTATACATTAATGTGAACCTCTTAACATTTTCACCA 745
QY 541 GAAATTTGCCATCGAAGTGCATGAGCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 600
DB 746 GAAATTTGCCATCGAAGTGCATGAGCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 805
QY 601 TTTTATACATTAATTAATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 660
DB 806 TCTTATACATTAATTAATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 865
QY 661 AATTAAGTATATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 720
DB 866 AATTAAGTATATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 925
QY 721 CTTTATACATTAATTAATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 780
DB 926 CTTTATACATTAATTAATGAGAACCCCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 985
QY 781 GACCATCTTTTATGTCCGAAATCTTCAACAAAGTAAAGATTAGAACTGTGGAA 840
DB 986 GACCATCTTTTATGTCCGAAATCTTCAACAAAGTAAAGATTAGAACTGTGGAA 1045
QY 841 GACTCAATTTGATAGTGGGATGCAATTTCTAATGATTAAGCTTTTCAAGTATCC 900
DB 1046 GACTCAATTTGATAGTGGGATGCAATTTCTAATGATTAAGCTTTTCAAGTATCC 1105
QY 901 AGTATAGTGTGATTTATTTTGAACAAAGACCTTTGATGTTGTCAGCCACTCCCAAT 960
DB 1106 AGTATAGTGTGATTTATTTTGAACAAAGACCTTTGATGTTGTCAGCCACTCCCAAT 1162
QY 961 AAAATGATGATTTTCAAGAAATTAAGTTTCACTGATTTTCTTCAAGGATGGA 1020
DB 1163 AAAATGATGATTTTCAAGAAATTAAGTTTCACTGATTTTCTTCAAGGATGGA 1219
QY 1021 AACAGTTTATTAATCTAGTGGGAAAT-----CAAGAAACAGTAAATGAGAAAGGGA 1074
DB 1220 AACAGTTTATTAATCTAGTGGGAAAT-----CAAGAAACAGTAAATGAGAAAGGGA 1279
QY 1075 AGATTAATTAAGTGAAGAAAGGCAATTTCTGATGTTTCTTCAAGGATGGA 1134
DB 1280 AGATTAATTAAGTGAAGAAAGGCAATTTCTGATGTTTCTTCAAGGATGGA 1339
QY 1135 TGTGATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1194
DB 1340 TGTGATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1396
QY 1195 TGTGATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1254
DB 1397 TGTGATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1453
QY 1255 AGTATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1314
DB 1454 AGTATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1512
QY 1315 AGTATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1360
DB 1513 AGTATGATCTGCACTTCTTAATAGTCAATGATCAAGCAAAACATTAATGAGAAAG 1558

```

RESULT 10  
 AAT08710  
 ID AAT08710 standard; cDNA; 1453 BP.  
 XX  
 XX AAT08710;  
 AC  
 AC  
 XX  
 DT 25-MAY-1996 (first entry)  
 XX  
 DE Sak serine-threonine kinase N-terminus-encoding cDNA.  
 XX  
 KM Sak; serine-threonine kinase; agonist; antagonist; proliferative disease;  
 KW cancer; tumour; antisense; transgenic animal; therapy; sb.

```

XX OS Mus musculus.
XX Key Location/Qualifiers
XX CDS 206..1453
XX FT /tag=a
XX FT /note="T bases may also be U"
XX CA2150789-A.
XX 03-DEC-1995.
XX 01-JUN-1995; 95CA-02150789.
XX 02-JUN-1994; 94US-00252995.
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX Dennis JM, Heffernan M, Fode C;
XX WPI; 1996-129817/14.
XX P-PSDB; AAR92176.
XX Nucleic acid encoding Sak serine-threonine kinase - useful for
XX identifying modulators potentially useful in treatment or prevention of
XX proliferative disease.
XX Claim 4; Page 46-48; 73pp; English.
XX A cDNA sequence (AAT08710) encodes the N-terminal 416 amino acids
XX (AAR92176) of 2 isoforms, sak-a and sak-b, of a novel serine/threonine
XX kinase associated with mitotic and meiotic cell division; sak-a and sak-b
XX differ in their C-terminal sequences (see AAT08711 and AAT08712). cDNA
XX clones for sak-a and sak-b were isolated from a murine lymphoid cell
XX (D3JW25) cDNA library established in CHO cells by selection of wheatgerm
XX agglutinin-resistant clones. Sak-a and sak-b are probably alternatively
XX spliced forms of the gene. The gene can be used for prodn. of recombinant
XX SAK, as a probe in the diagnosis of proliferative disorders or, in
XX antisense form, may be used to treat such disorders
XX SQ Sequence 1453 BP; 426 A; 315 C; 357 G; 355 T; 0 U; 0 Other;
XX Query Match 31.7%; Score 924.2; DB 2; Length 1453;
XX Best Local Similarity 86.4%; Pred. No.3.7e-230;
XX Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;
XX QY 1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60
XX DB 206 ATGGCGCCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 265
XX QY 61 GGATCATTTGCTGTCTTACAGAGTGAATCATTCACATGCTTTGGAAGTTGCAATC 120
XX DB 266 GGATCATTTGCTGTCTTACAGAGTGAATCATTCACATGCTTTGGAAGTTGCAATC 325
XX QY 121 AAAATGATGATAGAGAAAGCATGTACAAAGCAGAGATGTCAGAGATCCAAAATGAG 180
XX DB 326 AAAATGATGATAGAGAAAGCATGTACAAAGCAGAGATGTCAGAGATCCAAAATGAG 385
XX QY 181 GTGAAATATCATTTGCAATTTGAAACATCTTCTATCTTGGAGCTTATTAATTTTGA 240
XX DB 386 GTGAAATATCATTTGCAATTTGAAACATCTTCTGCTTGGAGCTTATTAATTTTGA 445
XX QY 241 GATAGCAATTTATGTATCTGTATTAAGAAATGTCATTAAGAGAAATGAAAGAGTAT 300
XX DB 446 GATAGCAATTTATGTATCTGTATTAAGAAATGTCATTAAGAGAAATGAAAGAGTAT 505
XX QY 301 CTAAAGAAATGAGTAAACCTCTCTGAGAAATGAGAGCTGACATTCATGACAGATC 360
XX DB 506 CTGAAAGAAAGAGTAAAGCTTTCTCAAGAAAGGAGAGCTAGGACATTCATGACAGAT 565
XX QY 361 ATCAGAGAGATGTTGATCTTCAATTCATGATGATATCTACACCGGAGACTCACTTCT 420
XX DB 566 ATCAGAGAGATGTTGATCTTCAATTCATGATGATGATGACGGGAGCTCACTCTCT 625

```

```

QY 421 AACCTCTACTGACTGTATATGAAACATCAAGATTGCTGATTTTGGCTGGCAACTCA 480
DB 626 AACATCTTACTTACCGGAAATATGAAACATTAATAATTTGCTGACTTTGACTTACGACG 685
QY 481 CTGAAATATGCAATGAAAGCACTATACATTTATGGAATCTCTAATCAATTTTCA 540
DB 686 TTGAATATGCAATGAAAGCACTATACATCTGTGGGACTCTTATATATTTTCA 745
QY 541 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGCTGGCTGTATG 600
DB 746 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGCTGGCTGTATG 805
QY 601 TTTTATACATTTCTATGCGGAGACACACCTTGACACTGACAGATCAAGAACATTA 660
DB 806 TCTTATACATTTCTATGCGGAGACACACCTTTTGACACTGACAGATCAAGAACATTTG 865
QY 661 AATTAAGTATATGCGAGATTAAGAAATGCAATCTTTTGTCAATAGAGGCAAGAC 720
DB 866 AACAAAGTATGCTCTGCGACATTAAGAAATGCGAGCTTTTGTCAAGAGGCCAGAC 925
QY 721 CTATATCAGCAATTTCTGTGAAATCCAGCAATGCTTTTAAGTTCTTCAATTTG 780
DB 926 CTATATCAGCAATTTCTGTGAAATCCCTGAGATGCTTAAATCTTCTGTGTTG 985
QY 781 GACCAATCTTTATATGCGGAAATCTTCAAGAAATTAAGATTAGAACTGTGAA 840
DB 986 GACCAATCTTTATATGCGGAAATCTTCAAGAAATTAAGATTAGAACTGTGAA 1045
QY 841 GACTCAATTAATGAGGAGATGCGCAATTTCTACTGCAATTAAGCTTCTTCAAGTAC 900
DB 1046 GACTCAATTAATGAGGAGATGCGCAATTTCTACTGCAATTAAGCTTCTTCAAGTAC 1105
QY 901 AGTATAGTGTATGTTTATTTGACAAAGAAAGCTTTGATTTGTGAGGCACTTCCAAAT 960
DB 1106 AGTATAGTGTATGTTTATTTGACAAAGAAAGCTTTGATTTGTGAGGCACTTCCAAAT 1162
QY 961 AAAATGATGATTTTCCAAAGAAATTAAGTTCAAGTATTTTCTTCAAGAGATGA 1020
DB 1163 AAAATGATGATTTTCCAAAGAAATTAAGTTCAAGTATTTTCTTCAAGAGATGA 1219
QY 1021 AACAGTTTATATCTCACTGAGGGAAT-----CAAGAAACAGTAATAGTGAAGGGA 1074
DB 1220 AATTAATTTTATATCTCACTGAGGGAATCAAGAAAGTAAATAGTGAAGGGAAGGGA 1279
QY 1075 AAGTATATCAAGTGCAGAGAAAGGCAATTTCTGATACCTTGTGTAGGCTTATTC 1134
DB 1280 AAGTATATCAAGTGCAGAGAAAGGCAATTTCTGATACCTTGTGTAGGCTTATTC 1339
QY 1135 TGTGATGATCTGGGCACTTCTAATAGTCTCAAGCAAAACATATCAATGAAAGCA 1194
DB 1340 TGTGATGATCTGGGCACTTCTAATAGTCTCAAGCAAAACATATCAATGAAAGCA 1396
QY 1195 TGTCACTCAGCAAGAAATGCTTCAAGTGTCCAAAGATCA 1233
DB 1397 TGTCACTCAGTGAATATGCTTCAAGCTTAAGATCA 1435

```

RESULT 11  
 ACH34348  
 ID ACH34348 standard; cDNA; 484 BP.  
 XX ACH34348;  
 AC 13-OCT-2003 (first entry)  
 XX  
 DE Human endothelial cell cDNA #2481.  
 XX  
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 XX genome mapping; biodiversity; genetic disorder.  
 OS Homo sapiens.  
 XX

PN US2003073623-A1.  
XX 17-APR-2003.  
XX 30-JUL-2001; 2001US-00918995.  
XX 30-JUL-2001; 2001US-00918995.  
XX 30-JUL-2001; 2001US-00918995.  
XX (DRMA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
PI Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 21560; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridization). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPRO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
CC  
SQ Sequence 484 BP; 167 A; 77 C; 100 G; 137 T; 0 U; 3 Other;  
Query Match 14.6%; Score 426; DB 8; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.6e-100;  
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2026 GACAAATTACGAGAAATATCTGGGAGAAATATCAATATGCTTCAGAGTTTGTACAGCTT 2085  
DB 59 GACAAATTACGAGAAATATCTGGGAGAAATATCAATATGCTTCAGAGTTTGTACAGCTT 118  
QY 2086 GTAAATCTAAATCTCCCAAAATCACTATTTTACAAATATGCTAAATGCAATTTTGTATG 2145  
DB 119 GTAAATCTAAATCTCCCAAAATCACTATTTTACAAATATGCTAAATGCAATTTTGTATG 178  
QY 2146 GAGAAATCTCCGTCGTCGATTTTGAAGTTGTTTATGATGGGTTAAATAACAAA 2205  
DB 179 GAGAAATCTCCGTCGTCGATTTTGAAGTTGTTTATGATGGGTTAAATAACAAA 238  
QY 2206 ACAGAGATTTCAATCAGGTGATGAAAGACAGGAGAGTCTTACCTTTAAAGTGA 2265  
DB 239 ACAGAGATTTCAATCAGGTGATGAAAGACAGGAGAGTCTTACCTTTAAAGTGA 298  
QY 2266 AGTGAAGTTAATAGCTTGAAGAGAGATTAATAATGTAATGACCAATGCTAATGAGGT 2325  
DB 299 AGTGAAGTTAATAGCTTGAAGAGAGATTAATAATGTAATGACCAATGCTAATGAGGT 358  
QY 2326 CATGCTATTTTGTAGCACTGGAATCCATATTTTCAAGAGGAAAGAAACCTAGAGAT 2385  
DB 359 CATGCTATTTTGTAGCACTGGAATCCATATTTTCAAGAGGAAAGAAACCTAGAGAT 418

QY 2386 GCTCCCTTTTCCCAATATCATAGAGAAACCTGGTAGTACTAGTTCACTAAGGCC 2445  
DB 419 GCTCCCTTTTCCCAATATCATAGAGAAACCTGGTAGTACTAGTTCACTAAGGCC 478  
QY 2446 TTATCA 2451  
DB 479 TTATCA 484  
RESULT 12  
ABL36507  
ID ABL36507 standard; cDNA; 331 BP.  
XX  
AC ABL36507;  
XX  
XX 08-APR-2002 (first entry)  
XX  
XX Human colon tumour antigen polynucleotide SEQ ID NO:96.  
XX  
XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;  
XX colon tumour metastatic antigen; diagnosis; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200196388-A2.  
XX  
XX 20-DEC-2001.  
XX  
XX 08-JUN-2001; 2001WO-US018557.  
XX  
XX 09-JUN-2000; 2000US-0210899P.  
XX  
XX 20-FEB-2001; 2001US-0270216P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Harlocker SL, Secrist H;  
XX  
XX WPI; 2002-114514/15.  
XX  
XX Novel isolated colon tumor polynucleotide differentially expressed in  
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,  
PT useful for inhibiting development of cancer in patient.  
XX  
XX  
XX Claim 1; SEQ ID NO 96; 105pp; English.  
XX  
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (1)  
XX which were isolated from human colon tumour and colon metastatic tumour  
XX cDNA libraries. (1) have cytostatic activity and can be used in vaccine  
XX production. (1) can be used for stimulating and/or expanding T cells  
XX specific for a tumour protein on contact with the T cells. They are also  
XX useful for inhibiting the development of cancer in a patient. (1) can be  
XX used as probes or primers for nucleic acid hybridisation, for preparing  
XX mutant species primers, or primers for use in genetic constructions. (1)  
XX can be used in the diagnosis of a colon tumour  
XX  
SQ Sequence 331 BP; 120 A; 68 C; 65 G; 76 T; 0 U; 2 Other;  
Query Match 11.3%; Score 329.2; DB 6; Length 331;  
Best Local Similarity 99.4%; Pred. No. 2.2e-75;  
Matches 329; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1428 ACAACAGTGGTTTGGGAATCTGCAATTAATGCTCAATTTAAGAAAACTAATGATGA 1487  
DB 1 ACAACAGTGGTTTGGGAATCTGCAATTAATGCTCAATTTAAGAAAACTAATGATGA 60  
QY 1488 CAGCATCAGCCCAACCGGACTTCCAGGGCCATCCAGTTTGCAGAGAGACATCAAA 1547  
DB 61 CAGCATCAGCCCAACCGGACTTCCAGGGCCATCCAGTTTGCAGAGAGACATCAAA 120  
QY 1548 AATAGCTGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCAATTG 1607  
DB 121 AATAGCTGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCAATTG 180

QY 1608 TGTAAACAGCAAAATACATGAAATATATGACCTGACCTTACAGTAACCTGAGATAT 1667  
 DB 181 TGTAAACAGCAAAATACATGAAATATATGACCTGACCTTACAGTAACCTGAGATAT 240  
 QY 1668 CCAACAGAAATGTTTTGGCTCAGATCTCTTTCTGAAACAGAGAACTNAGGGGTAT 1727  
 DB 241 CCAACAGAAATGTTTTGGCTCAGATCTCTTTCTGAAACAGAGAACTNAGGGGTAT 300  
 QY 1728 GGAGCCACATGGGGTATCATGAAATGCTACA 1758  
 DB 301 GGAGCCACATGGGGTATCATGAAATGCTDNA 331

## RESULT 13

ADD33867  
 ID ADD33867 standard; DNA; 609 BP.

AC ADD33867;

DT 15-JAN-2004 (first entry)

DE Mouse mitochondrial DNA sequence SEQ ID NO:1640.

XX de; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KM mitochondrial disease; oxidative phosphorylation dysfunction;  
 KM oxidative stress; apoptosis; aging.

XX Mus musculus.

PN WO2003020220-A2.

PD 13-MAR-2003.

PF 30-AUG-2002; 2002WO-US027886.

PR 30-AUG-2001; 2001US-0316323P.

PR 31-AUG-2001; 2001CA-02356540.

PA (UYEM-) UNIV EMORY.

PI Wallace DC, Levy S, Kerstann K, Procaccio V;

DR MPI; 2003-300821/29.

PT Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.

PS Claim 2; SEQ ID NO 1640; 201pp; English.

XX The invention relates to a novel array comprising at least two isolated  
 CC nucleotide molecules, each molecule having a sequence capable of uniquely  
 CC hybridizing to a nucleic acid molecule which is an expression product of  
 CC a gene involved in mitochondrial biology. The array comprises two or more  
 CC isolated nucleic acid molecules or spots, each molecule having a sequence  
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array  
 CC of the invention is useful for determining an expression profile of a  
 CC mouse or human sample containing nucleic acid, by contacting the array  
 CC with the sample under conditions allowing selective hybridisation, and  
 CC measuring hybridisation of nucleic acid in the sample to the array to  
 CC produce an expression profile. The array is also useful for determining  
 CC an expression profile of a first labelled sample containing nucleic acid  
 CC relative to a second, differently labelled sample containing nucleic  
 CC acid. The second sample is a reference or a standard. An array is useful  
 CC for determining an expression profile diagnostic of an energy-metabolism-  
 CC related physiological condition. An array of the invention is useful for  
 CC determining mitochondrial biology gene expression profiles of organisms,  
 CC such as human, mice and closely related species, tissue and organs of  
 CC such organisms, which are useful for determining expression profiles  
 CC diagnostic of energy metabolism-related physiological conditions,  
 CC diagnosing such physiological conditions, identifying biochemical  
 CC pathways, genes, and mutations involved in such physiological conditions,  
 CC identifying therapeutic agents useful for preventing and/or treating such

CC physiological conditions, evaluating and/or monitoring the efficacy of  
 CC such therapies, and creating and identifying animal models of human  
 CC energy metabolism-related physiological conditions. An array is also  
 CC useful for defining expression signatures or profiles for mitochondrial  
 CC diseases, as well as distinguishing clinical disorders that result from  
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
 CC apoptosis and aging. An array of the invention contains probes of genes  
 CC not previously recognised to participate in mitochondrial biology. The  
 CC sequences shown in ADD33867-ADD35260 represent murine mitochondrial DNA  
 CC clones used to make the probes of the invention. Some sequences are not  
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
 CC 1906, 2408 and 2643.

SO Sequence 609 BP; 189 A; 121 C; 113 G; 186 T; 0 U; 0 Other;

Query Match 9.7%; Score 284; DB 9; Length 609;  
 Best Local Similarity 82.1%; Pred. No. 1.8e-63;  
 Matches 345; Conservative 0; Mismatches 60; Indels 15; Gaps 1;

QY 2494 GCATCTTCAAGAAATGCTCAATGATGCTGCTTCCCAACAGGACCAATCCTT 2553  
 DB 6 GCCTCAGCAAGCAAGCTGAGCTGAATGTCGCTTCCCAACAGCTCCAGAGACTC 65

QY 2554 AATCCCTCTATGCTTCAAAATGAAAGACCTGCTTTTACAATACAGCTTCTGAAACAGAC 2613  
 DB 66 AGTCCTTCAACGTGACAGTTGAAGACTTGGCCACACAGCGACTGCCACAGGAACAGGC 125

QY 2614 ATCTCTTAATATGCTTAAAGATGCTCTTCTTAATACGACCACTTTGAAATCTGTT 2673  
 DB 126 GTCTCTTCA-----GTCTCTTAATGCAACAGCTTTGAAATCTGTT 170

QY 2674 TTGTGAAATATGTTGGTGGCTACACAGTTAACTAGTGAAGCTGTGAGGTTCACTT 2733  
 DB 171 TTGTGAAATATGTTGGTGGCTACACAGTTAACTAGTGAAGCTGTGAGGTTCACTT 230

QY 2734 AATGATGGTCCAGTGTGTTGACAGGAGAGTCTTTATACATGATTATCTGACCA 2793  
 DB 231 AATGATGGTCCAGTGTGTTGACAGGAGAGTCTTTATACATGATTATCTGACCA 290

QY 2794 AATGTCAAACAACTATGATGAGAAATGAAATTTCCAGACATCATCAACAGAA 2853  
 DB 291 GATGTCAGACAACTATGATGAGAAATGAAATTTCCAGACATCATCAACAGAA 350

QY 2854 TTACAGTGTCTGCTTCCATCTTGTATGATTTTCTATCCGACTCTATTTTCATGA 2913  
 DB 351 TTACAGTGTCTTCTTCCATCTTGTATGATTTTCTATCCGACTCTATTTTCATGA 410

## RESULT 14

ABL12501  
 ID ABL12501 standard; cDNA; 2585 BP.

AC ABL12501;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 31985.

KW Drosophila; developmental biology; cell signalling; insecticide;

XX pharmacological; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX MPI; 2001-656860/75.  
 DR P-PSDB; ABB68398.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX

PS Claim 1; SEQ ID NO 31985; 21np + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB150511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/publicated\_sequences

CC Sequence 2585 BP; 764 A; 596 C; 569 G; 656 T; 0 U; 0 Other;

Query Match 9.7%; Score 282.2; DB 4; Length 2585;  
 Best Local Similarity 60.0%; Pred. No. 1e-62;  
 Matches 470; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 14 TCGGGGAGAGATGAGATTTTAAAGTTGGAATCGTGTGTAAGATCATTTGCTG 73  
 DB 104 TTGGAGAAACATTTGAGACTTGAAGTACAGACTTGCTGAGTAAAGGTGTTTGGCA 163  
 QY 74 GTGTCTACAGAGCTGATGCTCATTACACTGTTTGGAGTTGCAATCAAAATGATGATA 133  
 DB 164 CCGTCTACAGAGGCGCGTGCCTCACACTCACAGATGTGGCCATTAAGATGATGATA 223  
 QY 134 AGAAGCCATGTACAAAGAGAGATGTACAGAGATCCAAATGAGTGAATACATT 193  
 DB 224 AAAAATCAATCCAGGCACTGAGCTACATCAACGCTTGCCAGAGAAAGTCAACT 283  
 QY 194 GCCAATTTAAACATCTTCTATCTTGGAGCTTTATTAATATTTTGAAGATAGCAATTG 253  
 DB 284 CCGGCTGAGAGATCCCTCTGTGCTTCACTGATCACTTTCTTTCAGAGCCCACTATG 343  
 QY 254 TGTATCTGTATTAAGAAATGTGCCATTAATGAGAAATGAACAGTATCTTAAAGATAG 313  
 DB 344 TGTATTTGGTCTGAGCTGCGCCATTAATGAGGAGCTTATGCTATATGAACCACTCG 403  
 QY 314 TGAACCTTCTCAGAAATGAAGCTCGACCTTCATGACACAGATCAACAGGAGTGT 373  
 DB 404 CCAGACCTTCAAGAGAGAGAGAGCTCTTCAATCTGAAGAGGTGTGGCGGACTTT 463  
 QY 374 TGTATCTTCAATCTCATGTATTAATCAACCGGAGCTTCACTTTCTTAACCTCTACTGA 433  
 DB 464 TGTACTGTGACTCCCAACATCATATGCAACCGGAGATTTGCTGTCACCACTACTGCTTA 523  
 QY 434 CTCGTAAATTAAGAAATCAAGATTTGCTGATTTTGGCTGGCAATCAATGAAATGCCAC 493  
 DB 524 GAGGAGAAATGACCTCAAGATGAGCCGATTTTGAATGCGCAACCAATTAAGGACCTG 583  
 QY 494 ATGAAAGCACTAATATTAATGTGGAACCTCTTAATCAATTTCAACAGAAATTCCTACTC 553  
 DB 584 ATGAGCCGCAATATGACATATGTGTGAATCTCCGAATATTAATTTGCGTGAAGTGTATG 643  
 QY 554 GAAAGTCACATGGCTTGAATCTGATTTTGTCTCCCTGGCTGTATGTTTATACATTAC 613  
 DB 644 GAACTCTCAAGAGATGCGCGGAGAGCTGTGAGAGGTGATGATGATGTACACCTGCG 703  
 QY 614 TTATCGGAGACCACTCTTGCACATGACATCAAGAACATTAATTAATTAATATATAT 673  
 DB 704 TGGTTGAGCGCCGCGCTTGAACCGATGACAGTCAACCTTAAACAAATTTGTGA 763

QY 674 TGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCGCAAGACCTTATTCACAGT 733  
 DB 764 TGTCCAGTACATATGATCCGGCTCACTTGTCTTACAGAGCGCAGACCTGAATCAAT 823  
 QY 734 TACTGTGAAATATCAGAGATCGTTTAAGTCTGTCTTCAATTTGACCATCTTTTA 793  
 DB 824 TATTAAAAAGCTCCGCGACGAGCAATTAACCTGAGAGCTGTCTTGTCCATCCCTTTA 883  
 QY 794 TGT 796  
 DB 884 TGT 886

RESULT 15  
 ACH37505  
 ID ACH37505 standard; cDNA; 491 BP.  
 XX  
 AC ACH37505;  
 XX  
 DT 13-OCT-2003 (first entry)  
 XX  
 DE Human endotheelial cell cDNA #5638.

XX Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 XX  
 OS Homo sapiens.

XX US2003073623-A1.  
 XX 17-APR-2003.  
 PD 30-JUL-2001; 2001US-00918995.  
 PF 30-JUL-2001; 2001US-00918995.  
 PR 30-JUL-2001; 2001US-00918995.

PA (DRNA/) DEMANAC R T.  
 PA (LABA/) LABAT I.  
 PA (STRAC/) STACHE-CRAIN B.  
 PA (DICK/) DICKSON M C.  
 PA (JONE/) JONES L W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

DR MPI; 2003-615964/58.

PT New polynucleotide sequences obtained from various cDNA libraries, useful  
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
 PT mapping, in the recombinant production of protein, or in generating  
 PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 24717; 44np; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 CC determined by the technique of SBH (sequencing by hybridisation). Also  
 CC included is a purified polypeptide comprising a sequence corresponding to  
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
 CC are useful in diagnostics as expressed sequence tags (EST) for  
 CC identifying expressed genes or for physical mapping of the human genome,  
 CC in forensics, in assessing biodiversity, or in identifying human mutations  
 CC responsible for genetic disorders and other traits. The nucleotide  
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
 CC for chromosome and gene mapping, in the recombinant production of  
 CC protein, or in generating antisense DNA or RNA. The present polypeptide  
 CC is useful for generating antibodies specific for it. The present sequence  
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC [seqdata.uspto.gov/sequence.html?docID=20030073623](http://seqdata.uspto.gov/sequence.html?docID=20030073623)

XX Sequence 491 BP; 134 A; 104 C; 141 G; 106 T; 0 U; 6 Other;

Query Match 8.6%; Score 250; DB 8; Length 491;  
Best Local Similarity 100.0%; Pred. No. 1.2e-54;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGGCACTGCAATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA	60
Db	241	ATGGGCACTGCAATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA	300
Qy	61	GGATCATTTGCTGTGTCTTACAGAGCTGATTCACACTGTGTTGGAAGTTGCAATC	120
Db	301	GGATCATTTGCTGTGTCTTACAGAGCTGATTCACACTGTGTTGGAAGTTGCAATC	360
Qy	121	AAATGATAGATAGAGAGCCATGTACAAAGCAGAAATGTACAGAGATCCAAATGAG	180
Db	361	AAATGATAGATAGAGAGCCATGTACAAAGCAGAAATGTACAGAGATCCAAATGAG	420
Qy	181	GTGAAAATACATTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATACTATTTGAA	240
Db	421	GTGAAAATACATTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATACTATTTGAA	480
Qy	241	GATAGCAATT	250
Db	481	GATAGCAATT	490

Search completed: September 24, 2004, 15:16:48  
Job time : 1097 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 14:00:23 / Search time 7172 Seconds  
(without alignments)  
12128.909 Million cell updates/sec

Title: US-10-026-021-1  
Perfect score: 2913  
Sequence: 1 atcgagaccatgcacgcgggga.....cgactcccaatttcacatga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_estbam:\*  
3: em\_estcin:\*  
4: em\_estcmu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estcro:\*  
8: em\_esthc:\*  
9: gb\_estc1:\*  
10: gb\_estc2:\*  
11: gb\_estc3:\*  
12: gb\_estc4:\*  
13: gb\_estc5:\*  
14: gb\_estc6:\*  
15: em\_estctun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1069.2	36.7	1656	11	AK006459 Mus muscu
2	925.8	31.8	3402	11	AK045082 Mus muscu
3	826.6	28.4	879	9	AU130274 AU130274
4	779.2	26.7	1182	11	AK006827 Mus muscu

Result No.	Score	Query Match	Length	ID	Description
5	750.4	25.8	825	14	CD558072 AGENCOURT
6	735.6	25.3	3137	11	BC045337 Danio rer
7	685.4	23.5	937	11	BM903951 AGENCOURT
8	669.8	23.0	976	10	BF794900
9	623.2	21.4	1027	13	BF706811
10	620.6	21.3	860	9	AL705839
11	619.6	21.3	797	12	BG290478
12	616.4	21.2	691	12	BG255560
13	609	20.9	817	13	BU166217
14	571	19.6	826	13	BQ440730
15	567	19.5	878	14	CB196323
16	555	19.1	557	12	BM830522
17	554.2	19.0	870	10	BF215170
18	538.8	18.5	848	13	BU708651
19	537.8	18.5	874	10	BF245698
20	536.4	18.4	596	12	BM759840
21	534.6	18.4	585	9	AV716458
22	531.4	18.2	875	10	BF243227
23	526	18.1	770	12	BI829311
24	513.8	17.6	528	9	AL043645
25	512.4	17.6	660	14	CF914035
26	507.2	17.4	650	14	CA558867
27	500.4	17.2	643	14	CF915391
28	499	17.1	637	14	CF175127
29	499	17.1	638	14	CF914476
30	496.8	17.1	870	9	AI894043
31	494.8	17.0	559	9	AI435520
32	494.4	17.0	570	12	BM537120
33	493.4	16.9	625	14	CF174552
34	491.8	16.9	506	9	AL048822
35	491.2	16.9	939	10	BF970082
36	489	16.8	619	14	CD552690
37	485.4	16.7	587	10	AW117220
38	479.6	16.5	782	14	CD643986
39	475.8	16.3	753	12	BG917180
40	471.6	16.2	592	14	CD553656
41	470.8	16.2	552	10	BE683420
42	466.8	16.0	590	14	CF906371
43	465.2	16.0	584	14	CA560592
44	463.6	15.9	584	14	CD561189
45	463.4	15.9	659	13	BY762320

## ALIGNMENTS

RESULT 1  
AK006459  
LOCUS  
DEFINITION  
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1170028H20 product:serine/threonine kinase 18, full insert sequence.  
ACCESSION  
AK006459  
VERSION  
AK006459.1 GI:12839571  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL  
MEDLINE  
PUBMED  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL  
MEDLINE  
PUBMED  
11042159

	REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komoh,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazami,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsui,M., Yoneda,Y., Ishikawa,T., Ozawa,X., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and Hayashi,zaki,Y.
	JOURNAL MEDLINE PUMED REFERENCE AUTHORS		Riken Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
	TITLE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
	JOURNAL REFERENCE AUTHORS	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1656)
	TITLE		Adechi,J., Aizawa,K., Akabira,S., Akimura,T., Arai,A., Aono,H., Arkawa,T., Bono,H., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayasuna,T., Hiramoto,K., Hirooka,I., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Komoh,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Ota,C., Salto,H., Salto,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Murematen,M., and Hayashizaki,Y.
	JOURNAL TITLE		Submitted (10-JULY-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa Yokohama Institute; 1-7-22 Suehiro-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken-go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
	COMMENT		Please visit our web site ( <a href="http://genome.gsc.riken-go.jp/">http://genome.gsc.riken-go.jp/</a> ) for further details.
			cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'- GAGAGAAGACATCCCAAGTCGAAGCTCGTTTGTTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAAGAGGGCGGCCGCATTAAATTCGTGAGATTATTAATTAATCCCCCCCCC 3'] cDNA was cleaved with XhoI and SacI. Cloning sites, 5' end: XhoI, 3' end: SacI. Host: SOLR.
	FEATURES		location/Qualifiers
	source	1..1656	/organism="Mus musculus"
			/mol_type="mRNA"
			/strain="C57BL/6J"
			/db_xref="FANTOM_DB:1700028H20"
			/db_xref="MG:1910124"
			/db_xref="taxon:10090"
			/clone="1700028H20"
			/sex="male"
			/tissue_type="testis"
			/clone_lib="RIKEN full-length enriched mouse cDNA library"
			/dev stage="adult"
CDS		<1..1610	

```

/notes="unnamed protein product; putative
serine/threonine kinase 18 (MD)|GI:101783, GB|AK006459,
evidence: BLASTN, 100%, match=1656."
/codon_start=3
/protein_id="BAB34599.1"
/db_xref="GI:12839572"
/translation="KTYSVGRCHSVEMLSKPRSLDENQHSNHHCLGKTPFPFADQTP
PQEMVQDMFGNLOMNALHGETNEHHIVTSNPRFOQDPDIQDTLRNMTPTPRASKND
TSAAVHAVKQLSAMKYVSAHHKREVPQEPGALPHBQSKNMSMETLLOKQFTLSL
ITSLALHRLKPTROKTKKAAVSLIDSEEVCEVLELREQSKNMSMETLLOKQFTLSL
VYVNDNGRGFPPLADRPPLPTDNTSRVSFNDLPEKMYKQYVARSFOLVNSKPTTY
FTRAKCITLMEKNSGADPEVWFYFDGAKITHTENLIHILEKTGISTALKNENPSTLKE
EYVVMHANBGRICLSLESVISEEKKRSGSFPIIVGRKPRGNSKRALVAPV
DPSCKCKQASARSLVNSAAPLPQSPGLSPSTVYEGLTATHTATCTGVSSIPKSA
QILKSFEVKNVGNATOLTSQAVVWOFDPGSLQVLVQAGVSIYSTSPGQFTTRGVENK
LPEYIKRQICLCTSSILMFSPNPNQ"

```

**ORIGIN**

Query Match	36.7%	Score 1069.2;	DB 11,	Length 1656;
Best Local Similarity	81.9%;	Pred. No. 6.4e-246;		
Matches 1291; Conservative	0;	Mismatches 289;	Indels 27;	Gaps 4;

QY	1338	ACCTGATTAACATCAAGACATCTCCAAATCATCTTTGTCAGAGAAAACTCCTTTTCATT	1397
Db	62	ACTGAGATGAAATCAACAACAGTTCCAAATCATCATGTTCTAGAGAAAACTCCTTTTCATT	121
QY	1398	TGCAGACCCGACACCTCAGACCTGAAACCGTTCACACAGGTGTTGGGAATCTGCAAAATPAA	1457
Db	122	TGCAGAACAGACACTCAGATGGAATGTACAGAGGTGTTGGGAATCTGCAAAATGAA	181
QY	1458	TGCTCATTTTAAAGAAAACTACTGATATGACAGATCAGCCCAACCGGCACTTCCAGGG	1517
Db	182	TGCTCATTTTAAAGAGAACTAATGTAGCACACACGGTTAGGCCAACAAGATTTCCAGGA	241
QY	1518	CCATCCAAATTTTGCAGAAAGGACATCATAAAAATGCCCTGAGCTGATACAAAAGTCAAAAA	1577
Db	242	CTATCCAAATTTTGC---AGGACAGGTTACGMAAGCGTTGACCTGACACGAGAGCCACGAA	298
QY	1578	GAACCTGATCCTCTCTGATTAATGACATCTGTGTAAACACGAAATACCAATGAAATATAT	1637
Db	299	GAATCTGATATCTTCTGCCATGTCTTCAATGCTGTAAACACCTGAGTGCATGAAATACGT	358
QY	1638	GACTGCACTTCAAGTAAACCTGAGATATCCAACAAGAAATGTGTTTTGGCTCAAGATCC	1697
Db	359	GAGTGCACTACATTAAGCCTGAGGTCAATGCCACAGG-----GCCGGGCTTCAATCC	412
QY	1698	TCTTTTGGAAACAGGCAAGACTAGGGATATGAGCCACATGGGGTTATCAGATCTGAC	1757
Db	413	TCATTTTGGAAACAAAGCAAGATATGAAATATAGGATCGACCTGGGTTTACAGAAACCTAC	472
QY	1758	ATTTAAGAAGCTTATCATCTCCGTTGGTGTGTCACAGGTTTAAACCAATCAGACAGAAAC	1817
Db	473	CTTAAGAAGATTTATCATCTCTCTGATTTGCTCACAGATTAACCAATCAGACAGAAAC	532
QY	1818	CAAAAAGCTGTGGTGAACATCTTAATTCAGAGAGAGTGTGTGAGACTTGTAAAGA	1877
Db	533	CAAAAAGCTGTGGTGAACCTTGAATTCAGAGAGAGTGTGTGAGACTTGTGAAGA	592
QY	1878	GTATGCACTTCAGAAATATGTGAAAGATTTCTTCAGATATCTAGTATGAAATACAT	1937
Db	593	GTGTGCTCTGAAAGATATGTGAAAGAGTCTTCAGATATCAGATGAGGACATATGAT	652
QY	1938	CACATATTTATTCACAAATGTGTAGAGGTTTTCTCTGTGCTGATAGACACACCTCAGC	1997
Db	653	CACGTTTATTTACCCGACGATGGAAGAGCTTTCTCTTGTGTCGACAGACCTCCCTGTGC	712
QY	1998	TACTGACAAATCAGTATGATGACAGCTTTGACAAATTTACAGAAAAATATCTGGCGAAAAAT	2057
Db	713	TACTGACAAATCAGTATGATGACAGCTTTGACAAATCTACAGAAAAATATCTGGCGAAAAAT	772
QY	2058	TCAATATGCTTCCAGGTTGTATACAGCTTGTAAATCTAAATCTCCCAAATCATTATTT	2117
Db	773	TCAATATGCTTCCAGATTCATTCAGCTGTATAGATCTAAATCTCCCAAATCATTATTT	832

Oy		2118	TACCAAGATATCTAAATGCACTTTTGA TGAGAAATTCTCCTGGTGCGATTTTGAGCTTG	2177
Db		833	TACAGAATATCTAATATGTATTTTGATGAAAATTCCTCGTGCTGATTTCCGAGTTTG	892
Oy		2178	GTTTTATGATGGGTAAATAATACA CAATA CAGAAATTTCA TTCA GGTGATTTGAAAAC	2237
Db		893	GTTTTATGATGGAGCAAATAATACATTA AACTGAAAA TTTAATTCACATTA TTGAGAAAAC	952
Oy		2228	AGGGAAGCTTACACTTTAAAAAGTGA AGGAA GTTAA TAgCTTGAAAGAGGATTA	2297
Db		953	AGGGAATATCTATATATTTAAAAAATGAAAATGAAGTTACAGAGCTGAAAAGGAA GTTAA	1012
Oy		2298	AATGATATGAGACCATGCTAATGAGGGCATGCTATTTGTTAGACATGGAATCCATAT	2357
Db		1013	AGTATATATGAGCAATGCTAATGAGGGTCA CGGTATTTGCTTGTCATCTGGAATCTGTAT	1072
Oy		2358	TTCAAGAAAGGAAAAGAAAAGTGAAGTGCTCCCTTTTCCCAATATCATAGAGAAA	2417
Db		1073	CTGAGAGAGGAAAAGAAAGAACAGAGGGGTTCTTCAATTTCTCCCTATTAATCGTAGAAGAAA	1132
Oy		2418	ACCTGGTATGTA CTAGTTTCACTTAAGGCTTATACA CTCCTCTCTCTGTGGAATCAAAATTA	2477
Db		1133	ACCTGGTATATCTAGTTTCACTTAAAGCTTATACAGTCTC---CTGGAAGCCAAGCTG	1189
Oy		2478	CCCAACGAGAGATAGAGCATTTTCAACAGAAATGGTCATGCATAGTGTGCTCTTCCAA C	2537
Db		1190	CTGTAAAGGAGAGCAGAGCGCTCAGCAGACAGACTAGCTGAATATATGCCCTTTCCCA C	1249
Oy		2538	ACAGGCAACCAATCCTTAAATCCCTCTATGGTTACAAATGAAAGACCTTGGTCTTACAATAC	2597
Db		1250	ACAATCCCCAAGACCTCAGTCCCTTCCA CTGTACAGATTGAAGACTTGGCCACA CAGCGAC	1309
Oy		2598	AGCTTCTGGGA CAGACATCTCTTCTTAATAGTCTAAAAGATGTCTTCTTAATACAGCA	2657
Db		1310	TGCCACAGGAACAGAGCGCTCTCTCAA-----GCTCTCTTAATCTGCACA	1354
Oy		2658	ACCTTTGGAATCTGTTTTTGTGAAAAAATGTGTGGCTACACAGTTAACTAGTGAAGC	2717
Db		1355	GCTTTGAAAATCTGTTTTTGTGAAAAAATGTGTGGCTACACAGCTAACTAGCGAGC	1414
Oy		2718	TGTGTGGGTTCA GTTTAATGATGGGTCCAGTGGTTGTGACGAGAGTGTCTTAT	2777
Db		1415	TGTGTGGGTTCA GTTTAATGATGGGTCA CAGTTGGTGTCCAGGACGAGATATCTTCAT	1474
Oy		2778	CAGTTTAACTCACCAAAATGTGTCAAACA CTNAGGTAATGAGAAAAATGAAAATTAACGA	2837
Db		1475	CAGTTTAACTCACCAAGATGTGTCAACA CACTNAGGTAATGAGAAAAATGAAAATTAACCTGA	1534
Oy		2838	CTACATCAAA CAGAAATTAACAGTGTCTGTCTTCAATCTTTTGATGTTTCTTAATCCGAC	2897
Db		1535	ATACATCAAA CAGAAATTAACAGTGTCTTCTTCAATCTTTTGATGTTTCTTAATCCAAC	1594
Oy		2898	TCCTAATTTTCATTGA	2913
Db		1595	TCCTAATTTTCATTGA	1610
RESULT 2				
AK045082				
LOCUS				
DEFINITION	Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length	3402 bp	mRNA	linear
VERSION	AK045082.1			
KEYWORDS	mus musculus (house mouse)			
SOURCE	Htc; CAP trapper.			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.			

JOURNAL MEDLINE PUBMED	TITLE High-efficiency full-length cDNA cloning
REFERENCE	Meth. Enzymol. 303, 19-44 (1999)
AUTHORS	99279253 10349636
JOURNAL MEDLINE PUBMED	TITLE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
AUTHORS	Genome Res. 10 (10), 1617-1630 (2000) 11042159
JOURNAL MEDLINE PUBMED	TITLE Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, U., Nishikawa, K., Kitabuchi, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Ikawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
REFERENCE	Riken integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multicapillary sequencer
AUTHORS	Genome Res. 10 (11), 1757-1771 (2000) 11076861
JOURNAL MEDLINE PUBMED	TITLE The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
REFERENCE	Functional annotation of a full-length mouse cDNA collection
AUTHORS	Nature 409, 685-690 (2001)
JOURNAL MEDLINE PUBMED	TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
AUTHORS	Nature 420, 563-573 (2002)
JOURNAL MEDLINE PUBMED	TITLE Adachi, J., Aizawa, K., Akinura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hisokane, T., Hoti, F., Imoent, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kotoh, H., Kawai, J., Koijima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Satoh, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takahashi, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
REFERENCE	Direct Submision
AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL MEDLINE PUBMED	TITLE cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
REFERENCE	Tissues were provided by Dr. Tomohito Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Huroko Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge.
AUTHORS	Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/. Location/Qualifiers
JOURNAL MEDLINE PUBMED	TITLE 1. 3402
REFERENCE	/organism="Mus musculus" /mol_type="mRNA"



Matches	865;	Conservative	0;	Mismatches	10;	Indels	3;	Gaps	3;
QY	1081	ATTCAAGATGCGAAGAAAGGCGCATTTCGATACCTTCGTAGAGCTTATTCCTCTGAT							1140
Db	1	ATTCAAGATGCGAAGAAAGGCGCATTTCGATACCTTCGTAGAGCTTATTCCTCTGAT							60
QY	1141	AGATCTGGCACTTCTAATAGTCAAGTCTCAAGCAAAACATATACATGAGATGAC							1200
Db	61	AGATCTGGCACTTCTAATAGTCAAGTCTCAAGCAAAACATATACATGAGATGAC							120
QY	1201	TCAGCAGAAATCTTTCAGTGTCCAAAGATCAGAGAGGTGAAAAATGAGAGAGTAC							1260
Db	121	TCAGCAGAAATCTTTCAGTGTCCAAAGATCAGAGAGGTGAAAAATGAGAGAGTAC							180
QY	1261	TCACCCACAGACACATGCGCAATTTTAACTTTTAAAGAAAGCATTCAGTACT							1320
Db	181	TCACCCACAGACACATGCGCAATTTTAACTTTTAAAGAAAGCATTCAGTACT							240
QY	1321	TCTGATCTTTGAAAGACCTGATTAACATCAAGACACTCCCATCATCTTTGTCAGGA							1380
Db	241	TCTGATCTTTGAAAGACCTGATTAACATCAAGACACTCCCATCATCTTTGTCAGGA							300
QY	1381	AAACTCTCTTTTCAATTTGCAAGCCGACACCTCAGACTGAAACCGTACACAGTGT							1440
Db	301	AAACTCTCTTTTCAATTTGCAAGCCGACACCTCAGACTGAAACCGTACACAGTGT							360
QY	1441	GGGATCTTGCAGAAATTAATGCTCAATTAAGAAAACTAGTATGACAGCATCAGCCCA							1500
Db	361	GGGATCTTGCAGAAATTAATGCTCAATTAAGAAAACTAGTATGACAGCATCAGCCCA							420
QY	1501	AACCGGACCTTCAGGGCCATCAGATTTGCAAGAGACACATCAAAAATGCGTGAAT							1560
Db	421	AACCGGACCTTCAGGGCCATCAGATTTGCAAGAGACACATCAAAAATGCGTGAAT							480
QY	1561	GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATTCCTGTAAGACGAA							1620
Db	481	GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATTCCTGTAAGACGAA							540
QY	1621	AATACATGAATATATGATGCTGCACTTCAAGTAAACCTGAGATATCCACAGAAATG							1680
Db	541	AATACATGAATATATGATGCTGCACTTCAAGTAAACCTGAGATATCCACAGAAATG							600
QY	1681	GTTTTGGCTGAGATCTCTTCTGAAACAGAGCAAGACTAGGGGTATGAGGCACCATG							1740
Db	601	GTTTTGGCTGAGATCTCTTCTGAAACAGAGCAAGACTAGGGGTATGAGGCACCATG							660
QY	1741	GTTATCAGAAATCTGATTAAGAGCATTAATCTCCGTTGGTGTCTCAGAGTTTAA							1800
Db	661	GTTATCAGAAATCTGATTAAGAGCATTAATCTCCGTTGGTGTCTCAGAGTTTAA							720
QY	1801	CCATCAGACAGAAAAACCAAAAAGGCTGTGTGAGCATCTT-GATTCAAGAGAGGTGTG							1859
Db	721	CCATCAGACAGAAAAACCAAAAAGGCTGTGTGAGCATCTT-GATTCAAGAGAGGTGTG							780
QY	1860	TGTGAGAGCTT-GTAAAGAGATGATGATCTCAA-GAATATGTAAAGAAAGTCTTCAGATA							1917
Db	781	TGTGAGAGCTTGTAAAGAGATGATGATCTCAA-GAATATGTAAAGAAAGTCTTCAGATA							840
QY	1918	TCTAGTATGAGAAATGATGATCTATTTATTTATCCAA							1955
Db	841	TCTAGTATGAGAAATGATGATCTATTTATTTATCCAA							878

RESULT 4  
AK006827 1182 bp mRNA linear HTC 20-SEP-2003  
LOCUS AK006827  
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:11700057013 product:serine/threonine kinase 18, full insert sequence.  
ACCESSION AK006827  
VERSION AK006827.1 GI:12840099  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL	sequencing pipeline with 384 multicapillary sequencer
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	20530913
REFERENCE	11076861
REFERENCE	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
PUBMED	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 1182)
TITLE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hisakawa, T., Horii, F., Imochi, K., Ishii, Y., Itoh, M., Izawa, M., Kasekawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamoto, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submision
JOURNAL	Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTCTT-3'], cDNA was prepared by using trihaloase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adaptor of sequence [5'



cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."

Query Match 25.8%; Score 750.4; DB 14; Length 825;  
Best Local Similarity 99.0%; Pred. No. 2.5e-169;  
Matches 776; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

ORIGIN

Query 337 GCTCGACATTTGACGACGATCATGACGAGGATGTTGATTTCTCATTCATGATTA 396  
Db 4 GATAGACATTTGACGACGATCATTC-CAGGATGTTGATTTCTCATTCATGATTA 62  
397 CTACACCGGAGCCTGACATTTCTAATCTGATCTGATTAATGAAATCAAGATT 456  
Db 63 CTACACCGGAGCCTGACATTTCTAATCTGATCTGATTAATGAAATCAAGATT 122  
Query 457 GCTGATTTGGGCTGGCACTCAATGAAATGCCATGAAAGCACTATATATG 516  
Db 123 GCTGATTTGGGCTGGCACTCAATGAAATGCCATGAAAGCACTATATATG 182  
Query 517 GGAATCTCTAATCACTTACCAAAATTCGACATGACATGCTTGAATCT 576  
Db 183 GGAATCTCTAATCACTTACCAAAATTCGACATGACATGCTTGAATCT 242  
Query 577 GATGTTGGTCCCTGGCTGTAATGTTTATATATCTATATGCGGAGACCCCTCGAC 636  
Db 243 GATGTTGGTCCCTGGCTGTAATGTTTATATATCTATATGCGGAGACCCCTCGAC 302  
Query 637 ACTGACACAGTCMAAACAATTAAATGATGATGGCAGATTAATGATGATG 696  
Db 303 ACTGACACAGTCMAAACAATTAAATGATGATGGCAGATTAATGATGATG 362  
Query 697 TTTTGTCAATAGAGGCCAAGACCTTATACCAAGTACTTGTGAAATCCAGCAGAT 756  
Db 363 TTTTGTCAATAGAGGCCAAGACCTTATACCAAGTACTTGTGAAATCCAGCAGAT 422  
Query 757 CGTTTAAGTCTGCTCAGATTTGACCAATCCCTTATGTCGGAATCTTCAAGAAA 816  
Db 423 CGTTTAAGTCTGCTCAGATTTGACCAATCCCTTATGTCGGAATCTTCAAGAAA 482  
Query 817 AGTAAGATTTAGAACTGTGAAAGCTCAATGATGAGGAGTCCCAATTTCTACT 876  
Db 483 AGTAAGATTTAGAACTGTGAAAGCTCAATGATGAGGAGTCCCAATTTCTACT 542  
Query 877 GCAATTAACAGCTTCTTCCAGTACCAATGATGATGATTTTGAAGAAAGACCT 936  
Db 543 GCAATTAACAGCTTCTTCCAGTACCAATGATGATGATTTTGAAGAAAGACCT 602  
Query 937 TTGATTTGGTCAAGCACTCCCAATTAATGATGATTTCCAAAGATTAATGATCACT 996  
Db 603 TTGATTTGGTCAAGCACTCCCAATTAATGATGATTTCCAAAGATTAATGATCACT 662  
Query 997 GATTTTCTTCTTCAAGAGTGAAGACAGTTTATATCTAGTGGGAAATCAAGAAC 1056  
Db 663 GATTTTCTTCTTCAAGAGTGAAGACAGTTTATATCTAGTGGGAAATCAAGAAC 722  
Query 1057 AGTAATAGTGAAGGGGAAAGATTAATCAAGATGCAAGAAAGCCAC-ATTCTGATTA 1115  
Db 723 AGTAATAGTGAAGGGGAAAGATTAATCAAGATGCAAGAAAGCCACATTCTGATTA 782  
Query 1116 CCTT 1119  
Db 783 CCTT 786

RESULT 6  
BC045337 3137 bp mRNA linear HTC 13-FEB-2003  
LOCUS BC045337  
DEFINITION Danio rerio. Similar to serine/threonine kinase 18, clone  
ACCESSION BC045337  
VERSION BC045337.1 GI:28374204

KEYWORDS  
HTC.  
Danio rerio (zebrafish)

SOURCE  
Danio rerio

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami  
cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amadan@systemsbio.org](mailto:amadan@systemsbio.org)  
Anup Madan, Jessica Fahy, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRK Place: 101 Row: 9 Column: 24  
This clone has the following problem: frame shifted.

FEATURES  
source  
1..3137  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:2639484"  
/tissue\_type="Whole body, adult, (one male and one female, including unfertilized eggs)"  
/clone\_1ib="Sugano\_Kawakami\_zebrafish\_DRA"  
/lab\_host="DH10B"  
/note="Vector: pME18S-FL3"

ORIGIN

Query Match 25.3%; Score 735.6; DB 11; Length 3137;  
Best Local Similarity 57.5%; Pred. No. 1.1e-165;  
Matches 1663; Conservative 0; Mismatches 1129; Indels 100; Gaps 15;

Query 9 CTGCATCGGGGAGAAAGATCGAGATTTAAAGTTGAAATCTGCTTGTAAAGATCAT 68  
Db 21 CTCAATCGGTGACAAAGATTGAGGACTTCAAGGTTCTCACCTTGTGGTAAAGGCTCCTT 80  
Query 69 TGTGTGTCTTCAAGAGCTGATCCATTGACATGCTGTTGGAAGTGCATCAAAATGAT 128  
Db 81 TGTGTGTCTTCAAGAGCTGATTCGTAAACACAGGATTTGGAAGTGCATCAAAATGAT 140  
Query 129 AGATAAGAGCATGTACAAAGCAGAGTGTATCAGAGTCCAAATGAGGTGAAAT 188  
Db 141 TGAAGAGAGCATGTACAAAGCAGAGTGTATCAGAGTCCAAATGAGGTGAAAT 200  
Query 189 ACATTGCCAATTAAGATCTTCTATCTTGAAGCTTTAACTATTTGAAGATGCAA 248  
Db 201 TCAGTCCCGTTAAAGCATCCATCAGTTCTCGAGCTGTCAATTAATTTGAGGACAGTAA 260  
Query 249 TTAATGTATCTGTATTAAGAAATGTCATTAATGAGAAATGAAGATTTAAAGAA 308  
Db 261 TTAATGTATCTGTATTAAGAAATGTCATTAATGAGAAATGAAGATTTAAAGAA 319  
Query 309 TAGAGTGAAGCCTTCTCAGAAATGAAGTCACTTCAATGACCAAGATCATCAGAG 368  
Db 320 GAGGAAATTCCTTTACAGAGAGAGGAGGAGCATTTCAATGACCAAGATTTAAAGAG 379  
Query 369 GATGTTGATCTTCATTTCTCATGATTAATCAACCGGAGCTTCACTTCTTCACTCT 428  
Db 380 CATGCTGATCTTCATTTCTCATGATTAATCAACCGGAGCTTCACTTCTTCACTCT 439

Qy	429	ACTGATCGTAAATGAACTCAAGATTGCTGATTTTGGGCTGGCACTCACTGAAT	488
Db	440	GCTCACCACGAGATGAACATTTAAATGCGAGACTTTGGTTAGCCACCCCACTCAAGCT	499
Qy	489	GCCACATGAAAAGCATATACATTATGTGAACTCTTAATCACTTTACCAAGAAATGCG	548
Db	500	TCCACAGTGAAGACATTTACATATGTGGGCAACCCAACTACATTTCACCCGAAGTGGC	559
Qy	549	CATCGAAGTGCATATGAGCCCTGTAATGTAATGTTGGTCCCTGGGCTGTATGTTTATAC	608
Db	560	CACCGGAGCGCTCAACGCTGTAGAGTCAAGCGTGTGCTGCTGAGCTGATTTTACGC	619
Qy	609	ATTACTATGGGAGACCCCTTTCAGACTGACACAGTCAAGCAATTAATAAAT	668
Db	620	CTTTCTCAGAGACGACCGCCCTTTCAGACAGACCGCTTAACGACCTCTCATTAAGT	679
Qy	669	AGTATGGGAGATTAATGAATGCAATCTTTTGTCTCATATGAGGCCAAGACCTTATCA	728
Db	680	GGTGTGGAGAGATCAAAATGCCCATGACATATCAGCAGAGGCTCAGGATCTAAATCCA	739
Qy	729	CCAGTACTTCTGAGAAATCCAGACATCGTTTAAAGTCTGTCTTCAGTATTTGACCATCC	788
Db	740	GCAAGCTGTACAAAAGAACCCAGCCCTCAGACCGACCTTTCTGCAAGTGTGAGCAACC	799
Qy	789	TTTTATGTCCGAAATTTCTTCAACAAAAGTAAAGATTTAGAACTGTGGA--AGACTC	845
Db	800	GTTCATGATCAAAAGTGGGCCACCGCAGCAAAAGACTCCGGCAGCAGACAGGAGCTC	859
Qy	846	AATGTATAGTGGCATGCGCAATTTCTGCAATTAAGCTTTCAGATCACTGATCACTAT	905
Db	860	CATCGACAGGGAATGCTTACATATCAACAGC--TTCAAAAGCAACCAACACAGCAG	916
Qy	906	AAATGTAGTTTATTTGACAAAAGAACTTTTGTATGTGTGAGCCATCCCAATTAAT	965
Db	917	CAGCAGTCGGTTCAGAGGGGAAACAGGCGATGATTGGGACCGGTTGCCAGTCTGAT	976
Qy	966	GACTGATTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAAGATGGAACAG	1025
Db	977	GGACCAATCCCTAGCCATCCCATATTTCCACAGCTCAGTTTAAAGTGGGAGGA	1036
Qy	1026	TTTTTACTCAATGGGGAATCAAGAAACAGTAAATAGTGAAGGGGAAAGTAAATCA	1085
Db	1037	CTGGCAGCCAAATTCACAAAGATGATTTATCCAGATGGGCGTTGGAGAAATCCCGTGG	1096
Qy	1086	AAGTGAAGAAAGGCCACATTTCTGATACCTTGTGAGCTTATCTCTGATGATC	1145
Db	1097	TGCTGACAGCGGAAGCCCTCATCTTCGTACTAAGACGAGCTCATTTGCTGATGATC	1156
Qy	1146	TGCACTTCTAAATGATCAATCAAGCAAAACATATACATGGAACGATGTCACTCAGC	1205
Db	1157	TGCTGTGGGTACAGCCACATTCACAGAGGCAAG--GCTGGAGAAATGTCACTCAGA	1213
Qy	1206	AAAATGCTTTTCAATGTCCAAAGATCAGAGAGGTGAAATGAAGAGGATCACTCAGC	1265
Db	1214	AAGAGTGTCTTACGGGGCTGACGGTTATTTCCACAGACCTCCGCTATAGAAATGCAAC	1273
Qy	1266	CACAGACAACATGCCCAATTTTAACTTTTAAAGAAAAGACATCCAGTATGTTGG	1335
Db	1274	ACATGG-----AACTCCAAACATGACAGG	1288
Qy	1326	ATCTTTGAAGACCTGATTAACATCAAGCACTTCCATCATCTTTTCCAGAGAAAAC	1385
Db	1299	CTTCCTCTCCACCTGTAATAACACTGTATCTGGGATGTCTTTCTTACGT---C	1354
Qy	1386	TCTTTTTCATTTGACAGCCGACACTCAGACTGAAAACGATACACAGTGTGGGAA	1445
Db	1355	CATCATTTTCAAGACAACAATGCCGATTTCTCAAC--ACAGCCCTGTGTCAAGAA	1411
Qy	1446	TCTGCAATTAATGTCTATTTAAGAAAATCTAGAAATATATACAGCATCAGCCAAACG	1505
Db	1412	T-----GATGAGTGTTCAGAGGCGGCAAGCATGAGCGGTCAACAGCAGATGG	1462

Qy	1506	GAACTTCCAGGGCCATCCAGATTTTGCAAGAGACATCAAAAATGCTGTGATATAC	1565
Db	1463	GAGTTTCCATCTGAAACGAG-----GCCAATCGGAAACAAACCTCTGACGTGACAA	1516
Qy	1566	AAAAGTCAAAAAGAACTGATGCTTTGATTAATGCACTTTGTGTAACAGCAAAATAC	1625
Db	1517	GCCACGGGCTTCACTTCA--GCAGACCCCATCTCTTCCAGACAAACCCCGGCG	1575
Qy	1626	CATGAATATATGACTGCATTCACAGTAAACCTGAGATTAATCCAAAGAAATGTGTTT	1685
Db	1576	CCTGCAGGGAAGATCATTTGATCAGGACACATGTCAAGACTCA-----GCCATAT	1628
Qy	1686	TGGCTCAGATCTCTTCTTGAACAGAGAACTAGGGGTATGAGCCACATGGGTTA	1745
Db	1629	TCGACGCAAGTTTCCATGTCCACCGCTTCCAAAGGCAAGCAACAGAGAAAG	1688
Qy	1746	TCAGATGTATACATTAAGAGCATATACATCTCCGTTGTGTCACAGTTTAAACCAAT	1805
Db	1689	GACAAAGTGTCTTAAGAAAGCTTC--CTCCGCTGTGTCTGCGACATCCAACTTAT	1747
Qy	1806	CAGACAGAAAACAAAAGCTGTGTGAGCATCTTGAATTCAGAGAGGTGTGTGGA	1865
Db	1748	TCGACAGAAAGCTAAAAATGCTGTGTGAGCATTTTGGGCAATGAGAGGTGTATGGA	1807
Qy	1866	GCTTGTAAAGAGATGATCATCTCAAGAAATATGTAAGAAATCTTGAGATATCTATGA	1925
Db	1808	GCTGTGTAAGAGACAAAGGGCTCAGAGAGGTCAAAAGATCTTGTGATATCTGTGA	1867
Qy	1926	TGAAATACGATCATATTTATATCCAAATGTGTGATGAGTTTCTCTGTGATAG	1985
Db	1868	TGATCTATGTGTGACAGTGTACAGCTTAATGAAGAAAGGGTTCCTGTGTGATCA	1927
Qy	1986	ACCACTTCACTTCTGACAAATCATCATAGTACAGTTTGAACAATTTACCAAAAATA	2045
Db	1928	TCACCTTCACTTCACTTGAAGATCTCATCTGAGTTTGTATGATTTTACAGAAAATA	1987
Qy	2046	CTGGCAAAATATCAATATGCTTCCAGTTTGTACAGCTTGTAAATCTAATCTCCAA	2105
Db	1988	CTGGAAGAAATATCAGTACGCAAAAGTTTGTGACACTGTCAAACTTAAGATCCAA	2047
Qy	2106	AATCATTTATTTCAAGATATGCTAAATGCAATTTGATGAGAAATCTCTGTGTGTA	2165
Db	2048	AGTACCTTTATACCAAAATTTGCAAGTCAATGATGAGAACTCTCTAATCAGA	2107
Qy	2166	TTTGAAGTTTGGTTTATGATGGGTTAAATACCAAAACAGAGATTTCAATCAAGT	2225
Db	2108	CTGGAAGTGTGCTTCTATATGTGTGCAAAACCATTAAGACAGTGAAGGTGCGGT	2167
Qy	2226	GATTAAGAAACAGGGAAGCTTACACTTTTAAAGTGAAGTGAAGTTAATAGCTTGA	2285
Db	2168	GGTGAAGAAAGTGGGAAGTCAATACAGTGAAGGGTTACGTTGATGATGGGCTCA	2227
Qy	2286	AGAGGAATTAATATATATGACATGTCTAATGAGGTCATGTATTTGTTAGCACT	2345
Db	2228	TCTGAATGCAACTCTATATTAAGCTTTCTGTGAGAGGACATCAATGTGTCTT	2287
Qy	2346	GGAATCCATATTTCA---GAAGGAAAGGAAATCAGAGTGTCTCCCTTTTCCCAAT	2402
Db	2288	GGAAGTGAATCAGAGCTGAGAGAGAGGCGACGCTTAAGAACATCTTCTTCCAT	2347
Qy	2403	AATCATTAAGAAAAACCTGTATGATGATGTTTCACTTAAAGGCTTATCACTCTCTTC	2462
Db	2348	CACCATATGGAGAGAGCGCTG-----TGAACCGGATCTCTCAGCTCC	2389
Qy	2463	TGTGATTAATTAATTAACAGAGATGATGAGATTTTCAACAGAAAGGTATGATG	2522
Db	2390	CGCTCTTATCTTCAATCTTCTGCGCTCAGGGGTGACAGAGAGGTGACAGTGTG	2449
Qy	2523	TGTGCTTTTCCAAACAGGACCAATCTTATCTCTATGTTTACAAATGAAGACT	2582
Db	2450	TCATAGTCTCCGACAGATCTCAATC---ACTCATTAATGATTTCTATAGCGGGT	2506
Qy	2583	TGCTTTTAACATCAAGCTTCTGAAACAGACATCTTCTTAATAGTTAAAGATGTCT	2642

Db 2507 AGACTTGACACAGCCAGTGTGGCTTAAGCAGCTCTCCGGTCCACAAAGATGAACGAC 2566  
 Qy 2643 TCTTAATACAGCAACTTTTGAATCTGTTTGTGAAAAATGTTGGTGGCTACACA 2702  
 Db 2567 CGTTAACTCTGGAAAAAGTCTCAAGTCCATTTTGTGGCCACATGGAATGGGCTCTCA 2626  
 Qy 2703 GTTAACATGAGAGCTGTGTGGGTCAATTAAATGATGGGTCCAGTTGGTGTGAGAGC 2762  
 Db 2627 GCTAACCATGGGAGAGTCTGGGTTCAGTTTAATGATGGGTGGCACTGATGGTTAAAGT 2686  
 Qy 2763 AGAGTGTCTTATCATGTTAACTTACCAAAATGTCACAACTAAGTATGAGAAAA 2822  
 Db 2687 TGGCGTCTTGATCATTTTAACTCCCGGAGAGACATCAAGAGTATTAAGAGAA 2746  
 Qy 2823 TGAATAATACAGACTATCATCAAGAAATTAACAGTGTCTGTCTTCCATCTTTTGAAT 2882  
 Db 2747 CGAGAAATTACAGAGCTGGTGAAGAAATCCACTGTGTCCAGCATCTTGGGTCT 2806  
 Qy 2883 GTTTTCTAATCC 2894  
 Db 2807 TCTGGCCATCC 2818  
 RESULT 7  
 BM903951 937 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOUT 6695921 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5492458  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM903951  
 VERSION BM903951.1 GI:19353918  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 937)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM12114 row: j column: 11  
 High quality sequence stop: 637.  
 Location/Qualifiers  
 1..937  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5492458"  
 /isue\_type="telinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH\_MGC\_67"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."  
 ORIGIN  
 Query Match 23.5%; Score 685.4; DB 12; Length 937;  
 Best Local Similarity 95.7%; Pred. No. 1.1e-153;  
 Matches 738; Conservative 0; Mismatches 26; Indels 7; Gaps 3;  
 Qy 1544 CAAAAATCCTGAGCTGATACAAAAGTCAAAAGAACTCTGATGCTTGTGATATGAC 1603  
 Db 24 CAAAAATCCTGAGCTGATACAAAAGTCAAAAGAACTCTGATGCTTGTGATATGAC 83

Qy 1604 ATTCTGTAAACAGCAAAATATCAATGAATATATGACTCAGCTTCAAGTAAACCTGAGA 1663  
 Db 84 ATTCTGTAAACAGCAAAATATCAATGAATATATGACTCAGCTTCAAGTAAACCTGAGA 143  
 Qy 1664 TAATCCAAAGAAATGTGTTTTTGGTCCAGATCTCTTTCTGAACAGACAAAGTACGAG 1723  
 Db 144 TAATCCAAAGAAATGTGTTTTTGGTCCAGATCTCTTTCTGAACAGACAAAGTACGAG 203  
 Qy 1724 GTATGAGCCACATGGGTTATCAGAAATCTATTAAGAAAGCATTACATCTCCGTTGG 1783  
 Db 204 GTATGAGCCACATGGGTTATCAGAAATCTATTAAGAAAGCATTACATCTCCGTTGG 263  
 Qy 1784 TTGCTCAGAGTTAAACCAATCAAGACAGAAACCAAAAGGCTGTGGAGACATCTTG 1843  
 Db 264 TTGCTCAGAGTTAAACCAATCAAGACAGAAACCAAAAGGCTGTGGAGACATCTTG 323  
 Qy 1844 ATTCAGAGAGGTGTGTGTGAGCTTTGAAGAGATGCAATCTCAAGAAATGTGAAG 1903  
 Db 324 ATTCAGAGAGGTGTGTGTGAGCTTTGAAGAGATGCAATCTCAAGAAATGTGAAG 383  
 Qy 1904 AAGTTCTTCAATATCTATGATGAGAAATACATCACTATTTATATCCAAATGTGAT 1963  
 Db 384 AAGTTCTTCAATATCTATGATGAGAAATACATCACTATTTATATCCAAATGTGAT 442  
 Qy 1964 GAGGTTTCTCTGCTGATAGACCACTCACTACTGACAAATCAGTATGAGTACAGCT 2023  
 Db 443 GAGGTTTCTCTGCTGATAGACCACTCACTACTGACAAATCAGTATGAGTACAGCT 502  
 Qy 2024 TTGACAAATTTACAGAAAAATATCTGCGAAAAATATCAATATGCTTCAGGTTTGTAC 2083  
 Db 503 TTGACAAATTTACAGAAAAATATCTGCGAAAAATATCAATATGCTTCAGGTTTGTAC 562  
 Qy 2084 TTGTAAGATCTAATCTCCCAAAATCACTTATTTTCAAGATATGTAATGATTTTGA 2143  
 Db 563 TTGTAAGATCTAATCTCCCAAAATCACTTATTTTCAAGATATGTAATGATTTTGA 622  
 Qy 2144 TGAGAAATCTCTGCTGCTGATTTTGAAGTTTGTGTTATGATGGGTAAATACACA 2203  
 Db 623 TGAGAAATCTCTGCTGCTGATTTTGAAGTTTGTGTTATGATGGGTAAATACACC 682  
 Qy 2204 AAACGAAATTTCAATTCAGTATGTAAGAAAGACAGGAGCTTACAC---TTTAAAAA 2260  
 Db 683 AAACGAAATTTCTTTCCGGGAGTTGAAAGAAAGGAGCTTAACTTTTAAAAAG 742  
 Qy 2261 GTGAAAGTGAATTAATAC---TTGAAAGAGAGTAAATATGTAATG 2308  
 Db 743 GGAAGAAAGGAAATTAATACCTTGAAAGAGAGTAAATATGTTTGG 793  
 RESULT 8  
 BF794900 976 bp mRNA linear EST 12-JAN-2001  
 LOCUS 602256767P1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4339985 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF794900  
 VERSION BF794900.1 GI:12099954  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 976)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Louis Staudt, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be



10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

**Source**

```

1.1027
/organism="Mus musculus"
/mol_type="rna"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1700057013"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male testis"
/notes="site 1: XhoI; site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GAGGAGGAGAGGATCCAGACGCTCTTTTCTTTTCTTTTCTTTTAA 3'}. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence 15'
GAGGAGGAGAGGCGCCGCAATTATTCGAGTTAATTAATTATATCCCCCCCC
3'}. cDNA was cloned into the XhoI and BamHI sites. "

```

**ORIGIN**

Query Match	21.4%	Score 623.2;	DB 13;	Length 1027;
Best Local Similarity	89.6%	Pred. No. 9.6e-139;		
Matches 670;	Conservative	0;	Mismatches 78;	Indels 0;
				Gaps 0;

Oy	1	ATGGCGACCTGCAATCGGGGAGAAAGATCGAGATTTTAAAGTTGGAAATCTGCTGGTAA	60
Db	210	ATGGCGGGGTGCATCTGGGAGAGAGATCGAGGCTTTAAGTTGGAAATCTACTCGGTAAA	269
Oy	61	GGATCATTTGCTGGTGTCTACAGAGCTGAGTTCATTCACTGGTTTGAAGTTGCATTC	120
Db	270	GGATCATTTTGTGGTGTCTACAGAGCTGAGTTCATTCACTGGTTTGAAGTTGCATTC	329
Oy	121	AAAAATGATAGTAAGAAAGCCATGTACAAAGCAAGGATGTTACAGAGAGTCCAAAATAG	180
Db	330	AAAAATGATAGTAAGAAAGCCATGTACAAAGCTGAATGTTACAGAGAGTCCAAAATAG	389
Oy	181	GTGAATAATACATTTGCCAATTGAAACATCTCTTCTATCTTGAAGCTTTATTAATCTATTTGAA	240
Db	390	GTGAATAATACATTTGCCAAGTTGAAACACCCCTCTGTCTTGGAGCTTCTAATAATTACTTTGAA	449
Oy	241	GATAGCAATTATGTGTATCTGTGATTATGAATATGCCCATTAATGAGAAATGAACAGGAT	300
Db	450	GATACCAATTATGTCTACTCTGTATTTGAAATATGCCCAATATGGAATGAACAGATAT	509
Oy	301	CTAAGAATAGTGAAGAAACCTTCTCAGAAAAATGAAGCTGCACATTCATGCACACAGATC	360
Db	510	CTGAAGAACAGAAATGAAGCTTTCTCAGAAAAGGAAGTAGGCACTTCAATGCACACAGATT	569
Oy	361	ATCA CAGGAGATGTTGTATCTTCAATCTCATGATATACTCAACCGGAGCTTGCACATTTCT	420
Db	570	ATCACAGGAATGTTATATCTTCAATCTCATGGAATATATGCAACCGGAGCTTGCACATTTCT	629
Oy	421	AACCTCTACTGACTCTGTAATATGAACATCAAGATTGCTGATTTTGGGCTGCACACTCAA	480

Db	630	AACATCTTACTAGCGGAAATATGAACATAAAAATGTGTGACTTTGGACTAGCAACGAG	689
Qy	481	CTGAAAATGCGACATGAGAAAGCACTATATACATTAATGTGAAATCTCGTAATCAATTTGACCA	540
Db	690	TTGAATATGCGACATGAAAAGCACTATACACTCTGTGGAGCTCCCAATTAATTAATTCACCA	749
Qy	541	GAATATGCGACATGAGATGCACTATGCGCTTGATCTGTATGTTTGGTCCCTGGGCTGTATG	600
Db	750	GAATATGCACTCGAAGTGACATGTGACTTGAATCTGTATATTTGGTCAATGGGCTGTATG	809
Qy	601	TTTTATATCACTTATATCGGGAAGCAACCCCTGCACACTGACACAGTAAAGAACATTA	660
Db	810	TTTTATATGCTACTTATATGGAAGCAACCTTTGACATGACACAGTAAAGAACATTTG	869
Qy	661	AATTAAGTAGTATTTGGCAGATTATGAATGCCATCTTTTTGTCAATAGAGGCCAAGAC	720
Db	870	AACAAAGTAGTCTGGCAGATTATGAATGCCACCTTTTGTCAAGAGAGGCCAAGAC	929
Qy	721	CTTATTCACCGAGTACTCTGGTAGAAATC	748
Db	930	CTTATTCACCGAGTACTCTGGTAGAAATC	957

RESULT 10	
AL705839	
LOCUS	AL705839 860 bp mRNA linear EST 04-SEP-2001
DEFINITION	DKFZps86862238_r1 686 (synonym: nlcc3) Homo sapiens cDNA clone
	DKFZps86862238 5', mRNA sequence.
ACCESSION	AL705839
VERSION	AL705839.1 GI:19689194
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

**REFERENCE**  
Otte nwaelder, B., O dermaier, B., Mewes, W., Mewes, H. W., Well, B. and  
1 (bases 1 to 860)

TITLE  
JOURNAL  
COMMENT  
EST (Ottensweider, B., Obermaier, B., Mewes, H.W., Well, B. and  
Wiemann, S.)  
Unpublished (2001)  
Contact: MIPS  
were

This clone (DKR2p68B3238) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: [clone@rzpd.de](mailto:clone@rzpd.de).  
Ingoltsrieder Landstr. 1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: [s.wiemann@dkfz-heidelberg.de](mailto:s.wiemann@dkfz-heidelberg.de)  
sequenced by MedGenomix (Martinsried/Germany) within the CDNA sequencing consortium of the German Genome Project. No 31 sequence available.

## FEATURES

**Source**

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKR20686B2238"
/dev_stage="adult"
/lab_host="DH10B"
/lab_note_1b="686 (synonym: hicc3)"
/clone_vector="prip1EX2; site_1: sf11; site_2: sf11b
-dna_collection"

```

**ORIGIN**

Query Match	21.3%	Score 620.6;	DB 9;	Length 860;
Best Local Similarity	99.1%;	Pred. No. 3.9e-138;		
Matches 645; Conservative	0;	Mismatches 4;	Indels 2;	Gaps 2;

Oy 1 ATGCGACCTGCATCGGGGAGAGATTCAGGATTTTAAAGTTGGAATCTGCTGTAAA 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 207 ATGCGACCTGCATCGGGGAGAGATTCAGGATTTTAAAGTTGGAATCTGCTGTAAA 266



RNA sequence.  
 ACCESSION BG255560  
 VERSION BG255560.1 GI:12765376  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 691)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Plate: LLNL0302 row: 1 column: 14  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1..691  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4476061"  
 /issue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Query Match 21.2%; Score 616.4; DB 12; Length 691;  
 Best Local Similarity 98.4%; Pred. No. 3.9e-137;  
 Matches 665; Conservative 0; Mismatches 6; Indels 5; Gaps 4;  
 QY TTTTGTGTTGTCAGCCATCTCCCAATTAATGACTGTATTTCCAAAGATTAAGTTCAA 994  
 DB 1 TTTAGATTGTCAGCCATCTCCCAATTAATGACTGTATTTCCAAAGATTAAGTTCAA 60  
 QY CTGATTTTCTTCTTCAAGAGATGAAACAGTTTATCTCAGTGGGGAATCAAGAA 1054  
 DB 61 CTGATTTTCTTCTTCAAGAGATGAAACAGTTTATCTCAGTGGGGAATCAAGAA 120  
 QY 1055 CCAATTAATGAGGGAAGGGAAGATTAATCAAGATGCAAGAAAGCCCATTTGAT 1114  
 DB 121 CCAATTAATGAGGGAAGGGAAGATTAATCAAGATGCAAGAAAGCCCATTTGAT 180  
 QY 1115 ACCTTGTGAGGTTTATCTCTGATAGATCTTGGCACTTCTATATGATCTCAAGCA 1174  
 DB 181 ACCTTGTGAGGTTTATCTCTGATAGATCTTGGCACTTCTATATGATCTCAAGCA 240  
 QY 1175 AAACATATCAATGAAAGATGTCACCTCAGCAGAAATGCTTCAAGTCCAAAGATCAG 1234  
 DB 241 AAACATATCAATGAAAGATGTCACCTCAGCAGAAATGCTTCAAGTCCAAAGATCAG 300  
 QY 1235 GAGGAGGTGAAATGAAAGAGGTACTCAACCAAGCAACAATGCAATTTTAACT 1294  
 DB 301 GAGGAGGTGAAATGAAAGAGGTACTCAACCAAGCAACAATGCAATTTTAACT 360  
 QY TCTTTAA-AGAAAGACATCAGTATGTTCTTTGAAAGACCTGATTAACATCA 1353  
 DB 361 TCTTTAAAGAAAGACATCAGTATGTTCTTTGAAAGACCTGATTAACATCA 420  
 QY 1354 GCACCTCCCAATCATCTTTGTCTCAGGAAAAAATCTCTTT-TCCATTTGAGAACCCGAC 1412

DB 421 GCACCTCCCAATCATCTTTGTCTCAGGAAAAAATCTCTTTCTCATTTCAGACCCGACACC 480  
 QY 1413 TCAGACTGAACCGTAAACAGTGTGGGAAATGCAAAATATGCAATTTAAGAA 1472  
 DB 481 TCAGACTGAACCGTAAACAGTGTGGGAAATGCAAAATATGCAATTTAAGAA 540  
 QY 1473 AACTACTGAATATGACAGATGAGCCAAACCGGAGCTTCCAGGCGCATTCAGATTTGCA 1532  
 DB 541 AACTACTGAATATGACAGATGAGCCAAACCGGAGCTTCCAGGAGCATTCAGATTTGCA 600  
 QY 1533 GAGGACACATCAAA-AAATGCTTGCAGTATCAAAA-GTCAAAAGAACTCTGATGC 1589  
 DB 601 GAGGACACATCAAA-AAATGCTTGCAGTATCAAAAAGTCAAAAGAACTCTGATGC 660  
 QY 1590 TTCTGATTAATGACAT 1605  
 DB 661 TTCTGATTAATGACAT 676

RESULT 13  
 BU166217 817 bp mRNA linear EST 04-SRP-2002  
 LOCUS AGENCOURT\_7952535 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:6107893  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BU166217  
 VERSION BU166217.1 GI:22680169  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 817)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Plate: LLNL0351 row: e column: 14  
 High quality sequence stop: 562.  
 Location/Qualifiers  
 1..817  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6107893"  
 /issue\_type="from acute myelogenous leukemia"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_55"  
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgccctggcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGCATTTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGCCAGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN  
 Query Match 20.9%; Score 609; DB 13; Length 817;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;  
 Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2305	ATGACCACTACTAATAGGGGTCACTCGATTATTGTTTAAAGCATGAATCCATAATTTTCAGA	2364
Dd	3	ATGACCACTACTAATAGGGGTCACTCGATTATTGTTTAAAGCATGAATCCATAATTTTCAGA	62
QY	2365	GAGAAAAGAAAACTAGAGAGTGCTCCCTTTTTCCCAATTAATATAGAGAAAGAAACCCTGT	2424
Dd	63	GAGAAAAGAAAACTAGAGAGTGCTCCCTTTTTCCCAATTAATATAGAGAAAGAAACCCTGT	122
QY	2425	AGTACTAAGTTCACTTAAGGCCTTAATCAGCTCCCTCTCTGTGAGATTCAAATTAACCAAG	2484
Dd	123	AGTACTAAGTTCACTTAAGGCCTTAATCAGCTCCCTCTCTGTGAGATTCAAATTAACCAAG	182
QY	2485	AGAGATAGAGCATCTTTCAACAAGAAATGATGCAATAGCTGCTTTCCCAACAGGCA	2544
Dd	183	AGAGATAGAGCATCTTTCAACAAGAAATGATGCAATAGCTGCTTTCCCAACAGGCA	242
QY	2545	CCAATCTTAATCCCTCTATGTTTCAAAATGAAGACTTGTCTTAACAATCAGCTTCT	2604
Dd	243	CCAATCTTAATCCCTCTATGTTTCAAAATGAAGACTTGTCTTAACAATCAGCTTCT	302
QY	2605	GGAAAGACATCTCTTAATAGTCTTAAAGTATGTTCTTCTTAATCAGCAAACTTTTG	2664
Dd	303	GGAAAGACATCTCTTAATAGTCTTAAAGTATGTTCTTCTTAATCAGCAAACTTTTG	362
QY	2665	AAATCTGTTTTGTGAAAAATTTGTTGGGCTACACAGTTAACATGAGAGCTGTGTGG	2724
Dd	363	AAATCTGTTTTGTGAAAAATTTGTTGGGCTACACAGTTAACATGAGAGCTGTGTGG	422
QY	2725	GTTCACTTTAATGATGGGTGCCAGTTGGTTGTGACAGGAGAGTGTCTTATCAGTTAT	2784
Dd	423	GTTCACTTTAATGATGGGTGCCAGTTGGTTGTGACAGGAGAGTGTCTTATCAGTTAT	482
QY	2785	ACCTCACCAATATGTCACAAACAATCTGATGAGAGAAATGAAAAATTAACAGACTAATC	2844
Dd	483	ACCTCACCAATATGTCACAAACAATCTGATGAGAGAAATGAAAAATTAACAGACTAATC	542
QY	2845	AAACGAATTAACAGTGTCTGTCTTCACACTTTTGAATGTTTCTAAATCCGACTCTTAAT	2904
Dd	543	AAACGAATTAACAGTGTCTGTCTTCACACTTTTGAATGTTTCTAAATCCGACTCTTAAT	602
QY	2905	TTCATTGA 2913	
Dd	603	TTCATTGA 611	
<hr/>			
RESULT 14			
BQ440730 826 bp mRNA linear EST 24-MAY-2002			
LOCUS LOCUS			
DEFINITION AGENCOURT 7890535 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168199			
ACCESSION BQ440730			
VERSION BQ440730.1 GI:21179806			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE 1 (bases 1 to 826)			
AUTHORS NIH-MGC htp://mgc.nci.nih.gov/			
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D. Email: ggaabs@emil.nih.gov Tissue Procurement: ATCC/DCTD/DRP CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.B. Consortium/LINU at: http://image.lnl.gov Plate: LHAM3531 row: f column: 08 High quality sequence stop: 745. Location/Qualifiers			
FEATURES			

Source	1..826
ORIGIN	1..826
Query Match	19.6%; Score 571; DB 13; Length 826;
Best Local Similarity	100.0%; P-Id. No. 3.3e-126;
Matches 571; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Db	1 ATGGCGACCTGTCATCGGGAGAGATGAGATTTTAAAGTTGGAATTCGCTGGTAA 60
Qy	254 ATGGCGACCTGTCATCGGGAGAGATGAGATTTTAAAGTTGGAATTCGCTGGTAA 313
Db	61 GGATTCATTTGCTGGTGTCTACAGAGCTGATGATTCATTCACCTGGTTGGAAGTTGCATC 120
Qy	314 GGATTCATTTGCTGGTGTCTACAGAGCTGATGATTCATTCACCTGGTTGGAAGTTGCATC 373
Db	121 AAATATGATGATTAAGAAAGCCATGTCAAAAGGAGAAATGGTACAGAGATCCAAAATGAG 180
Qy	374 AAATATGATGATTAAGAAAGCCATGTCAAAAGGAGAAATGGTACAGAGATCCAAAATGAG 433
Db	181 GTGAAAATACATTTGCCAATTTGAAACATCTTCATCTTGAGCTTTATATATATTTTGA 240
Qy	434 GTGAAAATACATTTGCCAATTTGAAACATCTTCATCTTGAGCTTTATATATATTTTGA 493
Db	241 GATGACAAATTAATGTGATCTGGATTAAGAAATGGTCAATTAATGAGAAATGAAACAGTAT 300
Qy	494 GATGACAAATTAATGTGATCTGGATTAAGAAATGGTCAATTAATGAGAAATGAAACAGTAT 553
Db	301 CTAAAGAAATGAGTGAACCCCTTCTCAGAAAATGAAAGTCTGACACTTCATGACACAGATC 360
Qy	554 CTAAAGAAATGAGTGAACCCCTTCTCAGAAAATGAAAGTCTGACACTTCATGACACAGATC 613
Db	361 ATCCACAGGGATTTTGATTTTATCTCATGTATATCTACACCGGAGCTCACACTTCT 420
Qy	614 ATCCACAGGGATTTTGATTTTATCTCATGTATATCTACACCGGAGCTCACACTTCT 673
Db	421 AACCTCTACTGACTGTATATGAAACATCAAGATTTGATTTTGGGCTGGCACTCAA 480
Qy	674 AACCTCTACTGACTGTATATGAAACATCAAGATTTGATTTTGGGCTGGCACTCAA 733
Db	481 CTGAAAATGGCCACATGAAAAGCATATATCATTTATGTGAACTCTTAACTAATTTTCACCA 540
Qy	734 CTGAAAATGGCCACATGAAAAGCATATATCATTTATGTGAACTCTTAACTAATTTTCACCA 793
Db	541 GAAATTTGCCACTGAAAGTGCACATGTGACCTTG 571
Qy	794 GAAATTTGCCACTGAAAGTGCACATGTGACCTTG 824

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. David Rowe  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: NDA0040 row: 9 column: 23  
High quality sequence spot: 676.  
Location/Qualifiers  
1. 878

## FEATURES

Source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30138118"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH MGC 135"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;  
Normalized full-length enriched library from pooled mouse  
embryonic 1lb, maxilla and mandible, day 12.5, 13.5,  
14.5, and 15.5 (size selected for the 0.5-1 kb fragments)  
cloned directionally, priming method: Oligo-dT. cDNA  
enrichment: >1k bp. Average insert size 1.6k bp.  
Normalization (Cot value): 7.5 kb. Priming sequence:  
5'-GACTAGTTCAGATCGGACGGCCGCC(T)3' Tissue contributed  
by, David Rowe. Library constructed by Reagen, Invitrogen  
Corp."

## ORIGIN

Query Match 19.5%; Score 567; DB 14; Length 878;  
Best Local Similarity 82.3%; Pred. No. 3e-125;  
Matches 664; Conservative 0; Mismatches 140; Indels 3; Gaps 1;  
QY 1809 ACAGAAACCAAAAGGCTGTGTGAGCATCTGATTCAGAGAGAGGTGTGAGGCT 1868  
DB 1 ACAGAAACCAAAAGGCTGTGTGAGCATCTGATTCAGAGAGAGGTGTGAGGCT 60  
QY 1869 TGTAAAGAGATGATCTCAAGAAATATGTGAAGAAGTTCTTCAGATATCTAGTATG 1928  
DB 61 TGTGAAGAGGTGTGAGTGAAGATATGTGAAGAAGTCTTCAGATATCTAGTATG 120  
QY 1929 AAATACGATCACTTATTTATTCAGAAATGTGTGAGAGTTTCTCTTGTGATGACC 1988  
DB 121 GACTATGATCACTGTTATTTATCCGAAACGATGAAAGGCTTCTCTGTGACAGACC 180  
QY 1989 ACCCTCACTACTGACAAATAGTATGATGATGATGATGATGATGATGATGATGATG 2048  
DB 181 TCCCTGCTTACTGACAAATAGTATGATGATGATGATGATGATGATGATGATGATG 240  
QY 2049 GCGAAATATCAATATGCTTCCAGGTTTGTACAGCTTGTGAATCTTAATCTCCAAAT 2108  
DB 241 GCGGAAATATCAATATGCTTCCAGATTCATTCAGCTAGTAAATCTTAATCTCCAAAT 300  
QY 2109 CACTTATTTTACAAAGATATGCTTAATGATGATGATGATGATGATGATGATGATG 2168  
DB 301 CACTTATTTTACAAAGATATGCTTAATGATGATGATGATGATGATGATGATGATG 360  
QY 2169 TGAAGTTTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2228  
DB 361 CGAAGTTTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
QY 2229 TGAAGAAGAGAGAGTCTTACCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2288  
DB 421 TGAAGAAGAGAGAGTCTTATTAATTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 480  
QY 2289 GGAGATTAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2348  
DB 481 GGAAGTAAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 2349 ATCCATTAATTTGAG 2408  
DB 541 ATCTGTAATCTCTGAG 600  
QY 2409 AGGAG 2468  
DB 601 AGGAG 657  
QY 2469 TTCAATTATCCCAAGAGATAGAGCATCTTTCAACAGATGTGATCATAGTGTGCG 2528  
DB 658 CCAAGCTCTCTTAAGGAG 717  
QY 2529 TTCTCCACACAGAGACCAATCTTAATCTCTTAATGTTTACAAATGAAGAGCTTGT 2588  
DB 718 TTCCCAACAGAGATCCCAAGAGATCTTCACTGTGACAGTTGAAGAGAGAGAGAG 777  
QY 2589 TACAACTACAGCTTCTGGAACAGACAT 2615  
DB 778 CACAGGAGCTGCACAGGAACAGGCGT 804

Search completed: September 24, 2004, 20:22:01  
Job time : 7184 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 14:58:38 ; Search time 1305 Seconds  
(without alignments)  
11305.898 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913  
Sequence: 1 atcgagcactgcacgcggga.....cgacctatcttcattga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2913	100.0	2913	15	US-10-026-021-1 Sequence 1, Appl1
2	2911.4	99.9	2913	15	US-10-037-270-280 Sequence 280, App
3	2911.4	99.9	3937	16	US-10-117-722-280 Sequence 280, App
4	2908.2	99.8	3721	13	US-10-116-802-108 Sequence 108, App
5	2786.4	95.7	2836	13	US-10-425-114-26254 Sequence 26254, A
6	2686.8	92.2	3924	16	US-10-062-674-1868 Sequence 1868, App
7	426	14.6	484	10	US-09-918-995-21560 Sequence 21560, A
8	328	11.3	328	9	US-09-878-178-96 Sequence 96, Appl
9	328	11.3	328	14	US-10-046-935-96 Sequence 96, Appl
10	328	11.3	328	15	US-10-146-502-96 Sequence 96, Appl
11	276	9.5	445	16	US-10-062-674-914 Sequence 914, App
12	250	8.6	491	10	US-09-918-995-24717 Sequence 24717, A
13	216.8	7.4	302	12	US-10-152-319A-292 Sequence 292, App
14	154.2	5.3	381	16	US-10-305-720-649 Sequence 649, App

15	135.6	4.7	2972	13	US-10-342-887-1302	Sequence 1302, App
16	135.6	4.7	2972	13	US-10-260-708-24	Sequence 24, Appl
17	135.6	4.7	2972	13	US-10-172-118-1302	Sequence 1302, App
18	135.6	4.7	2972	15	US-10-269-909-72	Sequence 72, Appl
19	135.6	4.7	2972	16	US-10-159-563-102	Sequence 102, App
20	134.4	4.6	2770	10	US-09-769-970-8	Sequence 8, Appl
21	134.4	4.6	2783	15	US-10-024-238A-102	Sequence 102, App
22	134.4	4.6	2783	15	US-10-042-211A-102	Sequence 102, App
23	134.4	4.6	2783	17	US-10-617-217A-102	Sequence 102, App
24	134.4	4.6	2786	16	US-10-264-049-949	Sequence 949, App
25	134.4	4.6	3024	15	US-10-198-846-11031	Sequence 11031, A
26	133.4	4.6	2781	12	US-10-152-319A-1953	Sequence 1953, App
27	133.4	4.6	2781	16	US-10-191-803-222	Sequence 222, App
28	131.4	4.5	2169	15	US-10-108-580-1	Sequence 1, Appl
29	131.4	4.5	2169	15	US-10-204-041-15	Sequence 15, Appl
30	131.4	4.5	2169	16	US-10-305-720-1147	Sequence 1147, App
31	131.4	4.5	2169	17	US-10-304-116-4	Sequence 4, Appl
32	131.4	4.5	2169	17	US-10-620-052A-27	Sequence 27, Appl
33	129.8	4.5	2378	17	US-10-755-889-3	Sequence 3, Appl
34	128.2	4.4	2410	17	US-10-304-116-12	Sequence 12, Appl
35	124	4.3	2778	15	US-10-059-585-13	Sequence 13, Appl
36	122	4.2	1104	16	US-10-369-493-46146	Sequence 46146, A
37	121.8	4.2	2118	16	US-10-369-493-2551	Sequence 2551, A
38	118	4.1	2033	13	US-10-342-887-881	Sequence 881, App
39	118	4.1	2033	13	US-10-172-118-881	Sequence 881, App
40	115.2	4.0	1539	9	US-09-938-842A-162	Sequence 162, App
41	115.2	4.0	1539	11	US-10-429-849-3	Sequence 3, Appl
42	113.6	3.9	1815	13	US-10-429-849-3	Sequence 915, App
43	113.2	3.9	2070	16	US-10-264-049-915	Sequence 83, Appl
44	113.2	3.9	2146	14	US-10-071-766-83	Sequence 2, Appl
45	113.2	3.9	2198	9	US-09-012-135A-2	

## ALIGNMENTS

RESULT 1  
US-10-026-021-1  
; Sequence 1, Application US/10026021  
; Publication No. US20030027756A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoshi, Yasumichi  
; APPLICANT: Demo, Susan  
; APPLICANT: Denkins, Yonchu  
; APPLICANT: Rigol Pharmaceuticals, Inc.  
; TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
; FILE REFERENCE: 021044-001210US  
; CURRENT APPLICATION NUMBER: US/10/026,021  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/309,632  
; PRIOR FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2913  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2913)  
; OTHER INFORMATION: human SAK serine/threonine kinase  
US-10-026-021-1  
Query Match 100.0%; Score 2913; DB 15; Length 2913;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2913; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGGCACTGATCGGAGAGATTCGAGTTTAAAGTTGGAATCTGTTGTA 60  
DB 1 ATCGGCACTGATCGGAGAGATTCGAGTTTAAAGTTGGAATCTGTTGTA 60  
QY 61 GATCATTTGCTGCTGCTACAGAGCTGATTCATTCACACTGTTGGAAGTTGCAATC 120

Db 61 ||||| GATCATTTGCTGGGTGTACAGAGCTGATGCCATTGCACTGGTTGGAGTTGCATC 120  
Qy 121 ||||| AAAATGATGATTAAGAAAGCCATGACAAACGAGAAATGATACAGAGTCCAAATGAG 180  
Db 121 ||||| AAAATGATGATTAAGAAAGCCATGACAAACGAGAAATGATACAGAGTCCAAATGAG 180  
Qy 181 ||||| GTGAAATACATTTGCCAATTGAAACATCTCTCATCTTGGAGCTTTATTAATTTTGA 240  
Db 181 ||||| GTGAAATACATTTGCCAATTGAAACATCTCTCATCTTGGAGCTTTATTAATTTTGA 240  
Qy 241 ||||| GATGCAATTAATGATGATCTGATTTAGAAATGTCCTAATGAGAAATGAAACAGGTAT 300  
Db 241 ||||| GATGCAATTAATGATGATCTGATTTAGAAATGTCCTAATGAGAAATGAAACAGGTAT 300  
Qy 301 ||||| CTAAAGATAGAGTGAACCCCTTCTCAGAAAATGAGCACTTCATGACCGAGATC 360  
Db 301 ||||| CTAAAGATAGAGTGAACCCCTTCTCAGAAAATGAGCACTTCATGACCGAGATC 360  
Qy 361 ||||| ATCAAGAGGATGTTGATCTTCATTCATGATGATTAATACACCGGGACTTCACTTTCT 420  
Db 361 ||||| ATCAAGAGGATGTTGATCTTCATTCATGATGATTAATACACCGGGACTTCACTTTCT 420  
Qy 421 ||||| AACCTCTACTGACTCGTAATATGAACATGAAATGCTGATTTGGGCTGGCACTCAA 480  
Db 421 ||||| AACCTCTACTGACTCGTAATATGAACATGAAATGCTGATTTGGGCTGGCACTCAA 480  
Qy 481 ||||| CTGAAAAATGCCACATGAAAAACATATATATGAGAACTCTCAATTCATTTACCA 540  
Db 481 ||||| CTGAAAAATGCCACATGAAAAACATATATATGAGAACTCTCAATTCATTTACCA 540  
Qy 541 ||||| GAAATGGCACTGGAAGTGCACATGACCTTGCATGATGTTGGTCCCTGGGCTGATG 600  
Db 541 ||||| GAAATGGCACTGGAAGTGCACATGACCTTGCATGATGTTGGTCCCTGGGCTGATG 600  
Qy 601 ||||| TTTTATACATTAATTAATGAGAACCACTTCGACATGACAGTCAAGAACATTA 660  
Db 601 ||||| TTTTATACATTAATTAATGAGAACCACTTCGACATGACAGTCAAGAACATTA 660  
Qy 661 ||||| AATTAAGTATGATGAGCAATTAATGAAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
Db 661 ||||| AATTAAGTATGATGAGCAATTAATGAAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
Qy 721 ||||| CTATATCACAAGTACTTGTGTAAGAAATGCAAGATCGTTTAAAGTGTCTTCAGATATG 780  
Db 721 ||||| CTATATCACAAGTACTTGTGTAAGAAATGCAAGATCGTTTAAAGTGTCTTCAGATATG 780  
Qy 781 ||||| GACCATCTTTTATGTCCTGGAATTTCTTCAACAAAAAGTAAAGATTTAGAACTGTGAA 840  
Db 781 ||||| GACCATCTTTTATGTCCTGGAATTTCTTCAACAAAAAGTAAAGATTTAGAACTGTGAA 840  
Qy 841 ||||| GACTCAATGATAGTGGGAGTGCACAAATTTCTAATGCAATTAAGCTTTTCCAGTACC 900  
Db 841 ||||| GACTCAATGATAGTGGGAGTGCACAAATTTCTAATGCAATTAAGCTTTTCCAGTACC 900  
Qy 901 ||||| AGTATAGTGTAGTATTTTGAACAAAAGAGACTTTTGATTTGTCAGCCACTCCCAAT 960  
Db 901 ||||| AGTATAGTGTAGTATTTTGAACAAAAGAGACTTTTGATTTGTCAGCCACTCCCAAT 960  
Qy 961 ||||| AAAATGACCTGATTTCCAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020  
Db 961 ||||| AAAATGACCTGATTTCCAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020  
Qy 1021 ||||| AACAGTTTATTAATCACTGAGGGAATCAAGAAACAGTAAATAGGGAAGGGAAGGTA 1080  
Db 1021 ||||| AACAGTTTATTAATCACTGAGGGAATCAAGAAACAGTAAATAGGGAAGGGAAGGTA 1080  
Qy 1081 ||||| ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTCTGAGGTTATTTCTCTGAT 1140  
Db 1081 ||||| ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTCTGAGGTTATTTCTCTGAT 1140  
Qy 1141 ||||| AGATCTGGCACTTCTAATAGTCAATTCACAAAAACATTAATAGAAACGATGTAC 1200  
Db 1141 ||||| AGATCTGGCACTTCTAATAGTCAATTCACAAAAACATTAATAGAAACGATGTAC 1200

Db 1141 ||||| AGATCTGGCACTTCTAATAGTCAATTCACAAAAACATTAATAGAAACGATGTAC 1200  
Qy 1201 ||||| TCAGCAAAATGCTTTCACTGTCACAAAAGATCAGAGGAGTGAATAGAAAGGATAC 1260  
Db 1201 ||||| TCAGCAAAATGCTTTCACTGTCACAAAAGATCAGAGGAGTGAATAGAAAGGATAC 1260  
Qy 1261 ||||| TCACCCACAGACACATGCCCCAATTTTAACTTTTAAAGAAAGACATCCAGTAT 1320  
Db 1261 ||||| TCACCCACAGACACATGCCCCAATTTTAACTTTTAAAGAAAGACATCCAGTAT 1320  
Qy 1321 ||||| TCTGATCTTTTGAAGACCTGATTAACATCAAGACCTCCAAATCACTTTTCCAGAG 1380  
Db 1321 ||||| TCTGATCTTTTGAAGACCTGATTAACATCAAGACCTCCAAATCACTTTTCCAGAG 1380  
Qy 1381 ||||| AAAATCTCTTTTCCATTTTGCACACCCGACCTCAGACTGAACCCGTACAAAGTGT 1440  
Db 1381 ||||| AAAATCTCTTTTCCATTTTGCACACCCGACCTCAGACTGAACCCGTACAAAGTGT 1440  
Qy 1441 ||||| GGGAAATCTGCAAAATTAAGTCAATTTAAGAAAATCTAATATATGACATCAAGCCA 1500  
Db 1441 ||||| GGGAAATCTGCAAAATTAAGTCAATTTAAGAAAATCTAATATATGACATCAAGCCA 1500  
Qy 1501 ||||| AACCGGGACTTCCAGGGCCATTCAGATTGCAAGAGACACATCAAAAATGCTGAGCT 1560  
Db 1501 ||||| AACCGGGACTTCCAGGGCCATTCAGATTGCAAGAGACACATCAAAAATGCTGAGCT 1560  
Qy 1561 ||||| GATTCAAAATGTCAAAAGAACTGATGCTTCTGATATGCAATTTGTAAAACAGCAA 1620  
Db 1561 ||||| GATTCAAAATGTCAAAAGAACTGATGCTTCTGATATGCAATTTGTAAAACAGCAA 1620  
Qy 1621 ||||| AATACATGAATATATGATCTGCACTTCAAGTAACTGAGATTAATCCAAAGAAATG 1680  
Db 1621 ||||| AATACATGAATATATGATCTGCACTTCAAGTAACTGAGATTAATCCAAAGAAATG 1680  
Qy 1681 ||||| GTTTTGGCTGATCTCTTTCTGAAACAGAGCAAGCTAGGGGTATGAGCCACATGG 1740  
Db 1681 ||||| GTTTTGGCTGATCTCTTTCTGAAACAGAGCAAGCTAGGGGTATGAGCCACATGG 1740  
Qy 1741 ||||| GGTATACGAATGATATTAAGAAAGATTAATCTCCCTTGTGTTGCTCAGGTTAAA 1800  
Db 1741 ||||| GGTATACGAATGATATTAAGAAAGATTAATCTCCCTTGTGTTGCTCAGGTTAAA 1800  
Qy 1801 ||||| CCAATCAGACAGAAAACCAAAAAGCTGTGTGAGATTAATTTCAAGAGAGTGTGT 1860  
Db 1801 ||||| CCAATCAGACAGAAAACCAAAAAGCTGTGTGAGATTAATTTCAAGAGAGTGTGT 1860  
Qy 1861 ||||| GTGGAGCTTGTAAAGAGATGATCTCAAGAAATATGTGAAGAGTCTTCAGATATCT 1920  
Db 1861 ||||| GTGGAGCTTGTAAAGAGATGATCTCAAGAAATATGTGAAGAGTCTTCAGATATCT 1920  
Qy 1921 ||||| AGTATGAAATACGATCACTATTTATTAATCCAAATGATGATGATTTTCTCTGTGCT 1980  
Db 1921 ||||| AGTATGAAATACGATCACTATTTATTAATCCAAATGATGATGATTTTCTCTGTGCT 1980  
Qy 1981 ||||| GATGACCACTTCACTCACTGACAAATCAATGAGTGAAGCTTTGACAAATTTACAGAA 2040  
Db 1981 ||||| GATGACCACTTCACTCACTGACAAATCAATGAGTGAAGCTTTGACAAATTTACAGAA 2040  
Qy 2041 ||||| AAATACGAGGAAATATGATATGCTTCAGGTTTGTACAGTTGTAAGATTAATCT 2100  
Db 2041 ||||| AAATACGAGGAAATATGATATGCTTCAGGTTTGTACAGTTGTAAGATTAATCT 2100  
Qy 2101 ||||| CCCAAATACATTTATTAACAAGATATGCTAAATGCAATTTGATGAGAGATTTCTCTGCT 2160  
Db 2101 ||||| CCCAAATACATTTATTAACAAGATATGCTAAATGCAATTTGATGAGAGATTTCTCTGCT 2160  
Qy 2161 ||||| GCTGATTTGAGTTTGTATGATGAGGTTAAATACAAATACAGAAATTCAT 2220  
Db 2161 ||||| GCTGATTTGAGTTTGTATGATGAGGTTAAATACAAATACAGAAATTCAT 2220  
Qy 2221 ||||| CAGGTATTTGAAAAGACAGGGAAGTCTTAACATTTAAAGTGAAGGATTAATGCT 2280  
Db 2221 ||||| CAGGTATTTGAAAAGACAGGGAAGTCTTAACATTTAAAGTGAAGGATTAATGCT 2280

QY 2281 TTGAAGAGAGATATAAATGTATATGACCATGCTAATAGGGTCATGCTATTGTTA 2340  
DB 2281 TTGAAGAGAGATATAAATGTATATGACCATGCTAATAGGGTCATGCTATTGTTA 2340  
QY 2341 GCACTGGAATCCATATTTTCAGAAAGAGAAAGAACTAGAGTCTCCCTTTTCCCA 2400  
DB 2341 GCACTGGAATCCATATTTTCAGAAAGAGAAAGAACTAGAGTCTCCCTTTTCCCA 2400  
QY 2401 ATATCATATGAAAGAAACCTGTAGTACTAGTTCACTTAAGCCTTATCACTCTCT 2460  
DB 2401 ATATCATATGAAAGAAACCTGTAGTACTAGTTCACTTAAGCCTTATCACTCTCT 2460  
QY 2461 TCGTGTGATTCAAATTAACCAAGAGATAGAGCATTTTCAACGAATGTCATGCAAT 2520  
DB 2461 TCGTGTGATTCAAATTAACCAAGAGATAGAGCATTTTCAACGAATGTCATGCAAT 2520  
QY 2521 AGTGTCTGCTTCTCAACACAGGACCAATCTTAATCCCTTAATGTTACAAATGAAGA 2580  
DB 2521 AGTGTCTGCTTCTCAACACAGGACCAATCTTAATCCCTTAATGTTACAAATGAAGA 2580  
QY 2581 CTGGTCTTACAACTACAGCTTCTGGAACAGACATCTTCTAATAGTCTAAAGATTGT 2640  
DB 2581 CTGGTCTTACAACTACAGCTTCTGGAACAGACATCTTCTAATAGTCTAAAGATTGT 2640  
QY 2641 CTTCCTAATTCAGCAACCTTTGAAATCTGTTTGTGAAATGTTGGTGGCTACA 2700  
DB 2641 CTTCCTAATTCAGCAACCTTTGAAATCTGTTTGTGAAATGTTGGTGGCTACA 2700  
QY 2701 CAGTTAACTAGTGAAGTGTGTGAGTTCAAGTTAATGATGGTCCAGTTGGTGTGAC 2760  
DB 2701 CAGTTAACTAGTGAAGTGTGTGAGTTCAAGTTAATGATGGTCCAGTTGGTGTGAC 2760  
QY 2761 GCAAGAGTCTTCTATCAGTTATACCTCAACAAATGTCAAACATAGATGAGAA 2820  
DB 2761 GCAAGAGTCTTCTATCAGTTATACCTCAACAAATGTCAAACATAGATGAGAA 2820  
QY 2821 AATGAAATTAACCAACATCAATCAACGAATTCAGAGTGTCTTCCATCCCTTTT 2880  
DB 2821 AATGAAATTAACCAACATCAATCAACGAATTCAGAGTGTCTTCCATCCCTTTT 2880  
QY 2881 ATGTTTCTAATCCGACTCTAATTTTCATTTGA 2913  
DB 2881 ATGTTTCTAATCCGACTCTAATTTTCATTTGA 2913

RESULT 2  
US-10-037-270-280  
Sequence 280, Application US/10037270  
Publication No. US20030104529A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyun  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aifeng J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yungling  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: Tiliinghaer, John  
APPLICANT: Dmanac, Radoje T.  
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
FILE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/10/037,270  
CURRENT FILING DATE: 2002-01-04

PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1104  
SOFTWARE: pc FL\_genes Version 1.0  
SEQ ID NO 280  
LENGTH: 3937  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (334) .. (3246)  
US-10-037-270-280

Query Match 99.9%; Score 2911.4; DB 15; Length 3937;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA 60  
DB 334 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA 393  
QY 61 GGATCATTTGCTGTGTCTACAGAGCTGATCATTCACTGTTGGAAGTTGCAATC 120  
DB 394 GGATCATTTGCTGTGTCTACAGAGCTGATCATTCACTGTTGGAAGTTGCAATC 453  
QY 121 AAATATATGATTAAGAAAGCCATGTACAAAGCAAGAAATGCTACAGAGATCCAAATGAG 180  
DB 454 AAATATATGATTAAGAAAGCCATGTACAAAGCAAGAAATGCTACAGAGATCCAAATGAG 513  
QY 181 GTGAAATATCATTTGCCAATTTGAAACATCTCTATCTTGAGCTTAACTATTGAA 240  
DB 514 GTGAAATATCATTTGCCAATTTGAAACATCTCTATCTTGAGCTTAACTATTGAA 573  
QY 241 GATAGCAATTAATGTATCTGTATTAAGAAATGTCATATGAGAAATGAAACAGATAT 300  
DB 574 GATAGCAATTAATGTATCTGTATTAAGAAATGTCATATGAGAAATGAAACAGATAT 633  
QY 301 CTAAAGATAGAGTGAACCTCTTCAAAATGAAGCTGACATTTCACTGACCGAGATC 360  
DB 634 CTAAAGATAGAGTGAACCTCTTCAAAATGAAGCTGACATTTCACTGACCGAGATC 693  
QY 361 ATCACAGGATGTGTATCTGATCTCATGATTAATCAACCGGAGCTCACTTCT 420  
DB 694 ATCACAGGATGTGTATCTGATCTCATGATTAATCAACCGGAGCTCACTTCT 753  
QY 421 AACCTCTACTGACTGTATATGAACATCAAGATTGCTGATTTGGCTGCACTCAA 480  
DB 754 AACCTCTACTGACTGTATATGAACATCAAGATTGCTGATTTGGCTGCACTCAA 813  
QY 481 CTGAAATATCCACATGAAGAACATATATATGGAATCTCTAATCAATTTACCA 540  
DB 814 CTGAAATATCCACATGAAGAACATATATATGGAATCTCTAATCAATTTACCA 873  
QY 541 GAAATGCGACCTGGAAGTGCACATGGCTTGAATCGATTTGGCTTGGCTGTATG 600  
DB 874 GAAATGCGACCTGGAAGTGCACATGGCTTGAATCGATTTGGCTTGGCTGTATG 933  
QY 601 TTTTATACATTACTTATCGGAGACCACTTTCGACATGACAGACATCAAGAACATTTA 660  
DB 934 TTTTATACATTACTTATCGGAGACCACTTTCGACATGACAGACATCAAGAACATTTA 993  
QY 661 AATTAAGTATATGCGAGATTAATGAATGCCATCTTTTGTCAATTAAGCCAAAGAC 720  
DB 994 AATTAAGTATATGCGAGATTAATGAATGCCAACTTTTGTCAATTAAGCCAAAGAC 1053  
QY 721 CTATATCACAGTCTTGTGAAATCGAGAGATCGTTAAGTGTCTTCAAGATTTG 780  
DB 1054 CTATATCACAGTCTTGTGAAATCGAGAGATCGTTAAGTGTCTTCAAGATTTG 1113  
QY 781 GACCATCTTTTATGTCCGAATTTCTTCACAAAGTAAAGATTAGAACTGTGAA 840

Db	1114	GACCATCTTTTATGATGCCGAATTTCTTCAAAAAAGTAAAGATTTAGAAGCTGTGAA	1173
Qy	841	GACTCAATTGATAGTGGCAATGCCAATTTCTATCTGCAATTAAGCTTCTTCCAGTACC	900
Db	1174	GACTCAATTGATAGTGGCAATGCCAATTTCTATCTGCAATTAAGCTTCTTCCAGTACC	1233
Qy	901	AGTAAAGTGTAGTATTATTTGACAAAAGAACCTTTTGATGGTCAAGCACTCCCAAT	960
Db	1234	AGTAAAGTGTAGTATTATTTGACAAAAGAACCTTTTGATGGTCAAGCACTCCCAAT	1293
Qy	961	AAAATGCTGATTTTCCAAAGATAAAAGTTCACTGATTTTCTTCTTCAAGATGGA	1020
Db	1294	AAAATGCTGATTTTCCAAAGATAAAAGTTCACTGATTTTCTTCTTCAAGATGGA	1353
Qy	1021	AACGTTTTTATCTAGTGGGGAATCAAGAAACGTAATGTGGAGGGGAAGATG	1080
Db	1354	AACGTTTTTATCTAGTGGGGAATCAAGAAACGTAATGTGGAGGGGAAGATG	1413
Qy	1081	ATTCAAGATCGAAGAAAGGCCAATTCTGATACCTTGTGAGCTTAATTCCTGTAT	1140
Db	1414	ATTCAAGATCGAAGAAAGGCCAATTCTGATACCTTGTGAGCTTAATTCCTGTAT	1473
Qy	1141	AGATCGGCACTCTAATAGTCACTCAAGCAAAAACATTAATCATGGAACATGTCAAC	1200
Db	1474	AGATCGGCACTCTAATAGTCACTCAAGCAAAAACATTAATCATGGAACATGTCAAC	1533
Qy	1201	TCAGCAAAATGCTTCAAGTGTCCAAAAGATCAGAGAGGTGAATGAAGAGGTAC	1260
Db	1534	TCAGCAAAATGCTTCAAGTGTCCAAAAGATCAGAGAGGTGAATGAAGAGGTAC	1593
Qy	1261	TCACCAACAGAACAAATGCCAATTTTAACTTCTTAAAGAAAACATCACTAGTAT	1320
Db	1594	TCACCAACAGAACAAATGCCAATTTTAACTTCTTAAAGAAAACATCACTAGTAT	1653
Qy	1321	TCGTGATCTTTTGAAGACCTGATTAACAAGCACTGCAATATCTTGTCCAGGA	1380
Db	1654	TCGTGATCTTTTGAAGACCTGATTAACAAGCACTGCAATATCTTGTCCAGGA	1713
Qy	1381	AAAACTCTTTTCAATTTGCAAGCCGACACCTCAGACTGAAACCGTACAAAGTGT	1440
Db	1714	AAAACTCTTTTCAATTTGCAAGCCGACACCTCAGACTGAAACCGTACAAAGTGT	1773
Qy	1441	GGGAATCTGCAATTAATGCTCATTTAAGAAAACATCTGAATTTGACAGCATCAAGCC	1500
Db	1774	GGGAATCTGCAATTAATGCTCATTTAAGAAAACATCTGAATTTGACAGCATCAAGCC	1833
Qy	1501	AACCGGACTTCAGGGCCATCAGATTTGCAAGAGGACATCAAAAAATCCTGGAAT	1560
Db	1834	AACCGGACTTCAGGGCCATCAGATTTGCAAGAGGACATCAAAAAATCCTGGAAT	1893
Qy	1561	GATCAAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCAATTTCTGTAAAAAGCA	1620
Db	1894	GATCAAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCAATTTCTGTAAAAAGCA	1953
Qy	1621	AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAAT	1680
Db	1954	AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAAT	2013
Qy	1681	GTTTTTGGCTCAATCTCTTTCTGAAACAGACAGACTAGGGGTATGGAAGCACAAGG	1740
Db	2014	GTTTTTGGCTCAATCTCTTTCTGAAACAGACAGACTAGGGGTATGGAAGCACAAGG	2073
Qy	1741	GGTATTCGAATCGATCAATTAAGAGCATTAATCTCCGTTGGTGTCTCAAGGTTAAA	1800
Db	2074	GGTATTCGAATCGATCAATTAAGAGCATTAATCTCCGTTGGTGTCTCAAGGTTAAA	2133
Qy	1801	CCAAATCAGACGAAAAACAAAAGGCTGTGTGACATATCTGATTGACAGAGAGGTGT	1860
Db	2134	CCAAATCAGACGAAAAACAAAAGGCTGTGTGACATATCTGATTGACAGAGAGGTGT	2193
Qy	1861	GTGAGCTTTGTAAAGAGATGCACTCAAGAAATGTGAAAGAGTTCTTCAAGATATCT	1920
Db	2194	GTGAGCTTTGTAAAGAGATGCACTCAAGAAATGTGAAAGAGTTCTTCAAGATATCT	2253

---

Qy	1921	AGTATGAAATACATCACTAATTTATATTCAAAATGCTGTAGAGTTTCTCTGTCT	1980
Db	2254	AGTATGAAATACATCACTAATTTATATTCAAAATGCTGTAGAGTTTCTCTGTCT	2313
Qy	1981	GATAGACACCCCTCACTGACAAACATGACTAGTACAGCTTTGACAAATTTACAGAA	2040
Db	2314	GATAGACACCCCTCACTGACAAACATGACTAGTACAGCTTTGACAAATTTACAGAA	2373
Qy	2041	AAATACGCGAAAATATCAATATCTTCCAGTTTGTACAGCTTGTAAATCTAATCT	2100
Db	2374	AAATACGCGAAAATATCAATATCTTCCAGTTTGTACAGCTTGTAAATCTAATCT	2433
Qy	2101	CCCAAAATCACTAATTTTACAAAGATATGCTAAATGCAATTTTGAAGAAATTCCTGT	2160
Db	2434	CCCAAAATCACTAATTTTACAAAGATATGCTAAATGCAATTTTGAAGAAATTCCTGT	2493
Qy	2161	GCTGATTTTGAAGTTTGGTTTTATGATGGGGTAAATAACAACAAAGAGAAATTCAT	2220
Db	2494	GCTGATTTTGAAGTTTGGTTTTATGATGGGGTAAATAACAACAAAGAGATTCAT	2553
Qy	2221	CAGGTGATTTGAAAAGACAGGGAAGCTTACACTTTAAAAGTGAAGTTAATAGC	2280
Db	2554	CAGGTGATTTGAAAAGACAGGGAAGCTTACACTTTAAAAGTGAAGTTAATAGC	2613
Qy	2281	TTGAAAGAGAGATPAAAATGATATGGAACAATGCTATGAGGTCATCGATTTGTTA	2340
Db	2614	TTGAAAGAGAGATPAAAATGATATGGAACAATGCTATGAGGTCATCGATTTGTTA	2673
Qy	2341	GCACTGGAATCCATTAATTTCAAGAAAGGAAAGAAAACATAGAGTGTCTCTTTTCCA	2400
Db	2674	GCACTGGAATCCATTAATTTCAAGAAAGGAAAGAAAACATAGAGTGTCTCTTTTCCA	2733
Qy	2401	ATTAATCATAAGAAAGAAAACCTGTGATCTAGTCACTTAAGCCTTATCACTCTCT	2460
Db	2734	ATTAATCATAAGAAAGAAAACCTGTGATCTAGTCACTTAAGCCTTATCACTCTCT	2793
Qy	2461	TCGTGATTCAAATTTCCCAACGAGAGATGAGCATCTTTCACACAGATGGTATGAT	2520
Db	2794	TCGTGATTCAAATTTCCCAACGAGAGATGAGCATCTTTCACACAGATGGTATGAT	2853
Qy	2521	AGTGTGCTTCTCCACACAGGACCAATCTTAATCCTCTATGTTTAAATGAAGGA	2580
Db	2854	AGTGTGCTTCTCCACACAGGACCAATCTTAATCCTCTATGTTTAAATGAAGGA	2913
Qy	2581	CTTGATCTTAACAATCAAGCTTCTGGAACAGACATCTCTTAATATCTTAAAAATGT	2640
Db	2914	CTTGATCTTAACAATCAAGCTTCTGGAACAGACATCTCTTAATATCTTAAAAATGT	2973
Qy	2641	CTTCTAATTCAGACAACTTTTGAATCTGTTTTTGTGAAAAATGTTGGTGGCTACA	2700
Db	2974	CTTCTAATTCAGACAACTTTTGAATCTGTTTTTGTGAAAAATGTTGGTGGCTACA	3033
Qy	2701	CAGTTAATCTAGTGAAGCTGTGTGGTTCAGTTTAATGATGGTCCAGTTGTTGACG	2760
Db	3034	CAGTTAATCTAGTGAAGCTGTGTGGTTCAGTTTAATGATGGTCCAGTTGTTGACG	3093
Qy	2761	GCAGAGTGTCTTATCACTTAATCTCACTCAACAAATGTCAAACAACTAGATGAGAA	2820
Db	3094	GCAGAGTGTCTTATCACTTAATCTCACTCAACAAATGTCAAACAACTAGATGAGAA	3153
Qy	2821	AATGAAAAATTAACAGACTACATCAACAGAAATTAAGTGTCTCTTCACTCTTTTG	2880
Db	3154	AATGAAAAATTAACAGACTACATCAACAGAAATTAAGTGTCTCTTCACTCTTTTG	3213
Qy	2881	ATGTTTCTAATCGAATCTCTAATTTTCAATGA	2913
Db	3214	ATGTTTCTAATCGAATCTCTAATTTTCAATGA	3246

RESULT 3  
US-10-117-722-280  
; Sequence 280, Application US/10117722

Publication No. US20030219744A1  
 GENERAL INFORMATION:  
 APPLICANT: Tang, Y. Tom  
 APPLICANT: Liu, Chenghua  
 APPLICANT: Abundi, Vinod  
 APPLICANT: Zhang, Jie  
 APPLICANT: Drmanac, Radoje T.  
 TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and  
 FILE REFERENCE: 784CIP2BCIP  
 CURRENT APPLICATION NUMBER: US/10/117,722  
 PRIOR FILING DATE: 2002-04-04  
 PRIOR APPLICATION NUMBER: 09/520,312  
 PRIOR FILING DATE: 2000-07-19  
 PRIOR APPLICATION NUMBER: 09/552,317  
 PRIOR FILING DATE: 2000-04-25  
 PRIOR APPLICATION NUMBER: 09/488,725  
 PRIOR FILING DATE: 2000-01-21  
 NUMBER OF SEQ ID NOS: 1104  
 SOFTWARE: pl\_Fl\_genes Version 1.0  
 SEQ ID NO 280  
 LENGTH: 3937  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (334)..(3246)  
 US-10-117-722-280

Query Match 99.9%; Score 2911.4; DB 16; Length 3937;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCACTGCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
 DB 334 ATGGGCACTGCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTGTAA 393  
 QY 61 GGATCATTTGCTGCTGCTACAGAGCTGATCCATTCACACGCTTGGAGTTGCAATC 120  
 DB 394 GGATCATTTGCTGCTGCTACAGAGCTGATCCATTCACACGCTTGGAGTTGCAATC 453  
 QY 121 AAAATGATGATAGTAAGAAAGCAATGTACAAAGCAGATGTAACAGAGTCCAAAATGAG 180  
 DB 454 AAAATGATGATAGTAAGAAAGCAATGTACAAAGCAGATGTAACAGAGTCCAAAATGAG 513  
 QY 181 GTGAAAATACATTTGCCAATTGAAACATCTTCTATCTTGGAGCTTTATTAATTTTGA 240  
 DB 514 GTGAAAATACATTTGCCAATTGAAACATCTTCTATCTTGGAGCTTTATTAATTTTGA 573  
 QY 241 GATAGCAATTAATGTATCTGTATGTAATGAAGTGCATTAATGGAATGAACAGTAT 300  
 DB 574 GATAGCAATTAATGTATCTGTATGTAATGAAGTGCATTAATGGAATGAACAGTAT 633  
 QY 301 CTAAAGATAGAGTGAACCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 360  
 DB 634 CTAAAGATAGAGTGAACCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 693  
 QY 361 ATCAAGGAGTGTGTATCTTCAATTCATGTAATCAACCGGAGCTCACAATTTCT 420  
 DB 694 ATCAAGGAGTGTGTATCTTCAATTCATGTAATCAACCGGAGCTCACAATTTCT 753  
 QY 421 AACCTCTCTGATCTGTATTAATGAACATCAAGATGCTGATTTTGGGCTGGAACTCAA 480  
 DB 754 AACCTCTCTGATCTGTATTAATGAACATCAAGATGCTGATTTTGGGCTGGAACTCAA 813  
 QY 481 CTGAAATGCAATGAAAGCACTATATATGTAAGTCTTAACTTCACTTCAATTCACA 540  
 DB 814 CTGAAATGCAATGAAAGCACTATATATGTAAGTCTTAACTTCACTTCAATTCACA 873  
 QY 541 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600  
 DB 874 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 933

QY 601 TTTTATACATTAATCTATCGGAGACCACTTTCAGACTGACAGTCAAGACATTA 660  
 DB 934 TTTTATACATTAATCTATCGGAGACCACTTTCAGACTGACAGTCAAGACATTA 993  
 QY 661 AATTAAGTATGTTGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
 DB 994 AATTAAGTATGTTGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 1053  
 QY 721 CTTTATCACCAGTTCCTGTGAAATCCAGAGATCGTTTAACTGTCTTCAGATTTG 780  
 DB 1054 CTTTATCACCAGTTCCTGTGAAATCCAGAGATCGTTTAACTGTCTTCAGATTTG 1113  
 QY 781 GACATCTTTTATGTCGGAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840  
 DB 1114 GACATCTTTTATGTCGGAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 1173  
 QY 841 GACTCAATGATAGTGGGATGCAATTTCTACTGCAATTAACGCTTTCAGTATC 900  
 DB 1174 GACTCAATGATAGTGGGATGCAATTTCTACTGCAATTAACGCTTTCAGTATC 1233  
 QY 901 AGTATAGTGTAGTTTATTTGACAAAAGAACTTTGATTGGTCAGCCACTCCCAAT 960  
 DB 1234 AGTATAGTGTAGTTTATTTGACAAAAGAACTTTGATTGGTCAGCCACTCCCAAT 1293  
 QY 961 AAAATGATGATATTTCCAAAGAAATTAAGTTCAACTGATTTTCTTTCAGAGATGA 1020  
 DB 1294 AAAATGATGATATTTCCAAAGAAATTAAGTTCAACTGATTTTCTTTCAGAGATGA 1353  
 QY 1021 AACAGTTTATCTCAGTGGGAAATCAAGAAACAGTAATAGTGAAGGGAGAGTA 1080  
 DB 1354 AACAGTTTATCTCAGTGGGAAATCAAGAAACAGTAATAGTGAAGGGAGAGTA 1413  
 QY 1081 ATTCAAGATGCAAGAAAGGCACTCTGATCTCTGATCTGATGAGCTTATCTCTGAT 1140  
 DB 1414 ATTCAAGATGCAAGAAAGGCACTCTGATCTCTGATCTGATGAGCTTATCTCTGAT 1473  
 QY 1141 AGATCGCACTTCTATATAGTCACTCAAGCAAAACATATATACATGAGATGTCAC 1200  
 DB 1474 AGATCGCACTTCTATATAGTCACTCAAGCAAAACATATATACATGAGATGTCAC 1533  
 QY 1201 TCAGCAAAATGCTTCACTGCTCAAGATCAAGAGAGTGAATGAAGAGGTAC 1260  
 DB 1534 TCAGCAAAATGCTTCACTGCTCAAGATCAAGAGAGTGAATGAAGAGGTAC 1593  
 QY 1261 TCAGCAAGAAATGCTTCACTGCTCAAGATCAAGAGAGTGAATGAAGAGGTAC 1320  
 DB 1594 TCAGCAAGAAATGCTTCACTGCTCAAGATCAAGAGAGTGAATGAAGAGGTAC 1653  
 QY 1321 TCTGATCTTTTGAAGAACTGATTAACATCAAGCACTCTCAATCATCTTTGTCCAGA 1380  
 DB 1654 TCTGATCTTTTGAAGAACTGATTAACATCAAGCACTCTCAATCATCTTTGTCCAGA 1713  
 QY 1381 AAAATCTTTTCAATTTGCAAGCCGACACTGAAACCGTACAAAGTGTGTT 1440  
 DB 1714 AAAATCTTTTCAATTTGCAAGCCGACACTGAAACCGTACAAAGTGTGTT 1773  
 QY 1441 GGAATCTGCAATTAATGCTATTTAAGAAAACTAGTATTAAGACGATTCAGCCCA 1500  
 DB 1774 GGAATCTGCAATTAATGCTATTTAAGAAAACTAGTATTAAGACGATTCAGCCCA 1833  
 QY 1501 AACCGGACCTTCAGGGCATCAAGATTGCAAGAGGACACATCAAAAATATGCTGAGCT 1560  
 DB 1834 AACCGGACCTTCAGGGCATCAAGATTGCAAGAGGACACATCAAAAATATGCTGAGCT 1893  
 QY 1561 GATACAAAAGTCAAAAAGAACTGTAGCTTGTGAATGCACTTCTGTAAACAGCAA 1620  
 DB 1894 GATACAAAAGTCAAAAAGAACTGTAGCTTGTGAATGCACTTCTGTAAACAGCAA 1953  
 QY 1621 AATACCAATTAATTAATGCTGCACTTCAAGTAAACCTGAAATTAATCAAGAAATGT 1680  
 DB 1954 AATACCAATTAATTAATGCTGCACTTCAAGTAAACCTGAAATTAATCAAGAAATGT 2013  
 QY 1681 GTTTTGGCTCAGATCTCTTCTGACAGACAAAGCTAGGGGTATGAGGCCACCATGG 1740

Db 2014 GTTTTGGCTGATCTCTTTCTGAACAGCAAGATTAAGGGATATGAGCCTACCATGG 2073  
Qy 1741 GGTTCAGAAATCCGATCACTTAAGAAGCACTTATCTCCGTTGGTCTCACAGGTTAAA 1800  
Db 2074 GGTTCAGAAATCCGATCACTTAAGAAGCACTTATCTCCGTTGGTCTCACAGGTTAAA 2133  
Qy 1801 CCAATCAGACAGAAACCAAAAGGCTGTGTGAGCACTTGAATTCAGAGAGAGTGTGT 1860  
Db 2134 CCAATCAGACAGAAACCAAAAGGCTGTGTGAGCACTTGAATTCAGAGAGAGTGTGT 2193  
Qy 1861 GTGAGCTTGTAAAGAGATATGCACTTCAAGAAATATGTGAAGAGTCTTCAAGATATCT 1920  
Db 2194 GTGAGCTTGTAAAGAGATATGCACTTCAAGAAATATGTGAAGAGTCTTCAAGATATCT 2253  
Qy 1921 AGTATGGAATATGATCACTATTTATATCAAAATGTGTGAGAGTGTTCCTCTTGTCT 1980  
Db 2254 AGTATGGAATATGATCACTATTTATATCAAAATGTGTGAGAGTGTTCCTCTTGTCT 2313  
Qy 1981 GATGACCACTCTCACTTACGACATCATGATGATGATGATGATGATGATGATGATGAT 2040  
Db 2314 GATGACCACTCTCACTTACGACATCATGATGATGATGATGATGATGATGATGATGAT 2373  
Qy 2041 AATATCTGGGAAATATCAATATGCTTCCAGGTTTGTATACCTTGTATGATGATGATGAT 2100  
Db 2374 AATATCTGGGAAATATCAATATGCTTCCAGGTTTGTATACCTTGTATGATGATGATGAT 2433  
Qy 2101 CCAAAATCACTTATTTTACAGATATGCTTCAAGATGATGATGATGATGATGATGATGAT 2160  
Db 2434 CCAAAATCACTTATTTTACAGATATGCTTCAAGATGATGATGATGATGATGATGATGAT 2493  
Qy 2161 GCTGATTTTGAAGTTGGTTTATGATGGGTTAAATATACAAACAGAGATTTTCAAT 2220  
Db 2494 GCTGATTTTGAAGTTGGTTTATGATGGGTTAAATATACAAACAGAGATTTTCAAT 2553  
Qy 2221 CAGGTGATGAAAGACAGGAGAGCTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 2280  
Db 2554 CAGGTGATGAAAGACAGGAGAGCTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2613  
Qy 2281 TTGAAAGAGAGAGATTAATATATATGATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2614 TTGAAAGAGAGAGATTAATATATATGATGATGATGATGATGATGATGATGATGATGAT 2673  
Qy 2341 GCACTGAATCCATATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
Db 2674 GCACTGAATCCATATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2733  
Qy 2401 ATATATATGAG 2460  
Db 2734 ATATATATGAG 2793  
Qy 2461 TCTGTGAGATTTCAAAATTAACCAACAGAGATTAAGAGATTTTCAACAGAGATTTTCA 2520  
Db 2794 TCTGTGAGATTTCAAAATTAACCAACAGAGATTAAGAGATTTTCAACAGAGATTTTCA 2853  
Qy 2521 AGTGTGCTTCTTCAACAG 2580  
Db 2854 AGTGTGCTTCTTCAACAG 2913  
Qy 2581 CTTGTGCTTCAACAG 2640  
Db 2914 CTTGTGCTTCAACAG 2973  
Qy 2641 CTTCTTAATGAGCACTTCTTGAATCTGTGTTTGTGAAGAAATGTTGTGTGGCTTACA 2700  
Db 2974 CTTCTTAATGAGCACTTCTTGAATCTGTGTTTGTGAAGAAATGTTGTGTGGCTTACA 3033  
Qy 2701 CAGTTTAATGAG 2760  
Db 3034 CAGTTTAATGAG 3093  
Qy 2761 GCAG 2820

Db 3094 GCAG 3153  
Qy 2821 AATGAAATATTTCCAGATCACTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 3154 AATGAAATATTTCCAGATCACTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3213  
Qy 2881 AGTGTGCTTCAACAG 2913  
Db 3214 AGTGTGCTTCAACAG 3246

RESULT 4  
US-10-116-802-108  
/ Sequence 108, Application US/10116802  
/ Publication No. US20030065157A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Amy Lasek  
/ TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER  
/ FILE REFERENCE: PA-0045 US  
/ CURRENT APPLICATION NUMBER: US/10/116,802  
/ PRIOR FILING DATE: 2002-04-04  
/ PRIOR APPLICATION NUMBER: 60/281,593  
/ NUMBER OF SEQ ID NOS: 519  
/ SOFTWARE: PERL Program  
/ SEQ ID NO 108  
/ LENGTH: 3721  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc feature  
/ OTHER INFORMATION: Incyte ID No: 492750CB1  
US-10-116-802-108

Query Match 99.8%; Score 2908.2; DB 13; Length 3721;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2910; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGACCTGCAATCCGGGAGAGAGATGAGAGATTTTAAAGTTGGAATCTGTTGGTAA 60  
Db 288 ATGGGACCTGCAATCCGGGAGAGAGATGAGAGATTTTAAAGTTGGAATCTGTTGGTAA 347  
Qy 61 GGATCATTTGCTGTGTCTCAAGAGTGAAGTCCATTCACCTGTTGGAAGTTGGAATC 120  
Db 348 GGATCATTTGCTGTGTCTCAAGAGTGAAGTCCATTCACCTGTTGGAAGTTGGAATC 407  
Qy 121 AAAATGATATGAAG 180  
Db 408 AAAATGATATGAAG 467  
Qy 181 GTGAAATATCAATGCAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
Db 468 GTGAAATATCAATGCAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 527  
Qy 241 GATGCAATATGATGTATGTGATTTAGAAATGTGCATATGAGAGAGAGAGAGAGAGAG 300  
Db 528 GATGCAATATGATGTATGTGATTTAGAAATGTGCATATGAGAGAGAGAGAGAGAGAG 587  
Qy 301 CTTAAAGATGAG 360  
Db 588 CTTAAAGATGAG 647  
Qy 361 ATCAAGAGAGATGTATATCTTCAATTCATGATGATATCAACAGGAGAGAGAGAGAGAG 420  
Db 648 ATCAAGAGAGATGTATATCTTCAATTCATGATGATATCAACAGGAGAGAGAGAGAGAG 707  
Qy 421 AACCTCCTATGAG 480  
Db 708 AACCTCCTATGAG 767  
Qy 481 CTGAAATATGAG 540  
Db 768 CTGAAATATGAG 827

QY 541 GAAATGGCACTGGAAGTGCACATGGCTTGATCTGATGTTGGTCCCTGGGCTGATG 600  
Db 828 GAAATGGCACTGGAAGTGCACATGGCTTGATCTGATGTTGGTCCCTGGGCTGATG 887  
QY 601 TTTTATACATTACTTATCGGAGACCACTTGCACATGACAGTCAAGAACATTA 660  
Db 888 TTTTATACATTACTTATCGGAGACCACTTGCACATGACAGTCAAGAACATTA 947  
QY 661 AATAAGTATGATGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
Db 948 AATAAGTATGATGGCAGATTATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 1007  
QY 721 CTTTATCACCAGTACTTGTGTAAATCCGACAGATCGTTAAGTCTGTCTTCAATG 780  
Db 1008 CTTTATCACCAGTACTTGTGTAAATCCGACAGATCGTTAAGTCTGTCTTCAATG 1067  
QY 781 GACCATCTTTTATGTCGGAATTTCTTCAACAAAAGTAAAGATTAGAGCTGTGAA 840  
Db 1068 GACCATCTTTTATGTCGGAATTTCTTCAACAAAAGTAAAGATTAGAGCTGTGAA 1127  
QY 841 GACTCAATTGATAGTGGGCAATGCAATTTCTAATGCAATTAACGCTTCTTCAAGTAC 900  
Db 1128 GACTCAATTGATAGTGGGCAATGCAATTTCTAATGCAATTAACGCTTCTTCAAGTAC 1187  
QY 901 AGTATAGTGTATGTTATTTTGA CAAGAAGACTTTGATGTTGTCAGACCACTCCAAAT 960  
Db 1188 AGTATAGTGTATGTTATTTTGA CAAGAAGACTTTGATGTTGTCAGACCACTCCAAAT 1247  
QY 961 AAAATAGTGTATTTCCAAAGATTAAGTTCAATGATTTTCTTCTTCAAGAGATGA 1020  
Db 1248 AAAATAGTGTATTTCCAAAGATTAAGTTCAATGATTTTCTTCTTCAAGAGATGA 1307  
QY 1021 AACAGTTTATTACTCAATGGGAAATCAAGAAACCAATATAGTGAAGGGAGAGTA 1080  
Db 1308 AACAGTTTATTACTCAATGGGAAATCAAGAAACCAATATAGTGAAGGGAGAGTA 1367  
QY 1081 ATTCAGATGCGAAGAAAGGCGCAATCTGCACTCTTCTGAGCTTATCTCTGAT 1140  
Db 1368 ATTCAGATGCGAAGAAAGGCGCAATCTGCACTCTTCTGAGCTTATCTCTGAT 1427  
QY 1141 AGATCTGGCACTCTTATAGTCACTCAAGCAAAAACATATCAATGAGACATGTCAC 1200  
Db 1428 AGATCTGGCACTCTTATAGTCACTCAAGCAAAAACATATCAATGAGACATGTCAC 1487  
QY 1201 TCAGCAGAAATGCTTCAAGTCCAAAGATCAGAGAGGTGAATGAAGAGGTAC 1260  
Db 1488 TCAGCAGAAATGCTTCAAGTCCAAAGATCAGAGAGGTGAATGAAGAGGTAC 1547  
QY 1261 TCACCCACAGACAAATGCGCAATTTTAACTTCTTAAAGAAAAGACATCAATAGT 1320  
Db 1548 TCACCCACAGACAAATGCGCAATTTTAACTTCTTAAAGAAAAGACATCAATAGT 1607  
QY 1321 TCTGATCTTTTGAAGACCTGATTAACAATCAAGACCTCCATCATCTTTTCCAGGA 1380  
Db 1608 TCTGATCTTTTGAAGACCTGATTAACAATCAAGACCTCCATCATCTTTTCCAGGA 1667  
QY 1381 AAAAATCTTTTTCATTTGCAAGCCGACACTCAAGCTGAAAACGTTACAAAGTGGTT 1440  
Db 1668 AAAAATCTTTTTCATTTGCAAGCCGACACTCAAGCTGAAAACGTTACAAAGTGGTT 1727  
QY 1441 GGGAACTGCAAAATTAATGCTCAATTTAAAGAAAATTAAGATTAAGACATCAAGCCA 1500  
Db 1728 GGGAACTGCAAAATTAATGCTCAATTTAAAGAAAATTAAGATTAAGACATCAAGCCA 1787  
QY 1501 AACCGGAACTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAATGCTGGAAT 1560  
Db 1788 AACCGGAACTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAATGCTGGAAT 1847  
QY 1561 GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATATGCAATCTGTAAACAGCA 1620  
Db 1848 GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATATGCAATCTGTAAACAGCA 1907

QY 1621 AATACATGAATAATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGATGT 1680  
Db 1908 AATACATGAATAATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGATGT 1967  
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACTAGGGGATAGAGCCCAATGG 1740  
Db 1968 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACTAGGGGATAGAGCCCAATGG 2027  
QY 1741 GGTATCAAAATGATCAATTAAGAACATTAATCACTCCGTTGGTTGCTCAGAGTTAAA 1800  
Db 2028 GGTATCAAAATGATCAATTAAGAACATTAATCACTCCGTTGGTTGCTCAGAGTTAAA 2087  
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGAATCAGAGAGGTGT 1860  
Db 2088 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGAATCAGAGAGGTGT 2147  
QY 1861 GTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTAAGAAAGTCTTCAAGATATCT 1920  
Db 2148 GTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTAAGAAAGTCTTCAAGATATCT 2207  
QY 1921 AGTATAGAAATACGATCACTATTTATCCAAAATGTTGATGAGGTTTCTCTTGGCT 1980  
Db 2208 AGTATAGAAATACGATCACTATTTATCCAAAATGTTGATGAGGTTTCTCTTGGCT 2267  
QY 1981 GATAGACCAACCTCACTCACTGACATCAAGTGAAGGTAAGCTTTGACAAATTAACAGAA 2040  
Db 2268 GATAGACCAACCTCACTCACTGACATCAAGTGAAGGTAAGCTTTGACAAATTAACAGAA 2327  
QY 2041 AAATACGCGAAATATCAATATGCTTCAGGTTTGTACAGTTGTAAAGATTAATCT 2100  
Db 2328 AAATACGCGAAATATCAATATGCTTCAGGTTTGTACAGTTGTAAAGATTAATCT 2387  
QY 2101 CCAAAAATCACTAATTTTCAAGATATGCTAAATGATTTTGA TGAAGATTTCTCTGGT 2160  
Db 2388 CCAAAAATCACTAATTTTCAAGATATGCTAAATGATTTTGA TGAAGATTTCTCTGGT 2447  
QY 2161 GCTGATTTGAGGTTGTTTATGATGAGGGTAAATACAAACAGAGATTTCAAT 2220  
Db 2448 GCTGATTTGAGGTTGTTTATGATGAGGGTAAATACAAACAGAGATTTCAAT 2507  
QY 2221 CAGGTATGAAAGACAGGAAATCTTACCTTAAAGTGAAGTAAATAGC 2280  
Db 2508 CAGGTATGAAAGACAGGAAATCTTACCTTAAAGTGAAGTAAATAGC 2567  
QY 2281 TTGAAAAGAGATTAATAATGATATGACATGCTAAATGAGGTCATGATATTTCTTA 2340  
Db 2568 TTGAAAAGAGATTAATAATGATATGACATGCTAAATGAGGTCATGATATTTCTTA 2627  
QY 2341 GCACTGAAATCCATATTTCAAGAGAGAAAGAAATAGAGAGTCTCCCTTTTCCCA 2400  
Db 2628 GCACTGAAATCCATATTTCAAGAGAGAAAGAAATAGAGAGTCTCCCTTTTCCCA 2687  
QY 2401 ATATATCAGAGAAAGAAACCTGTAGTACTAGTTCACTTAAGGCTTATCACTCTCTCT 2460  
Db 2688 ATATATCAGAGAAAGAAACCTGTAGTACTAGTTCACTTAAGGCTTATCACTCTCTCT 2747  
QY 2461 TCTGTGATTCAAATTAACCAACAGAGATTAAGACATTTTCAACAGAAATGTCAGCAT 2520  
Db 2748 TCTGTGATTCAAATTAACCAACAGAGATTAAGACATTTTCAACAGAAATGTCAGCAT 2807  
QY 2521 AGTGTGCTTCTTCAACAGAGCAACAAATCTTAAATCCCTTATGTTGTTCAATGAAGA 2580  
Db 2808 AGTGTGCTTCTTCAACAGAGCAACAAATCTTAAATCCCTTATGTTGTTCAATGAAGA 2867  
QY 2581 CTTGGTCTTAAACATCAAGCTTCTGGAACAGACATCTTCTAATAGTCTAAAGATGT 2640  
Db 2868 CTTGGTCTTAAACATCAAGCTTCTGGAACAGACATCTTCTAATAGTCTAAAGATGT 2927  
QY 2641 CTTCTTAAATCAAGCAAACTTTTGAATGCTTTTGTGAAAAAATGTTGGGCTACA 2700  
Db 2928 CTTCTTAAATCAAGCAAACTTTTGAATGCTTTTGTGAAAAAATGTTGGGCTACA 2987  
QY 2701 CAGTTAACTAGTGAAGTGTGAGGTTCAGTTTAAATGATGAGGTCAGAGTTGTTGTGAG 2760

Db 2388 CAGTAACTAGTGGAGCTGTGGCTTCACTTATGATGGGTCCAGTTGGTTGGCAG 3047  
Qy 2761 GCAGAGAGCTTCTTATCAGTTATACCTCCAAATGGTCAAAACATAGTATGAGAA 2820  
Db 3048 GCAGAGAGCTTCTTATCAGTTATACCTCCAAATGGTCAAAACATAGTATGAGAA 3107  
Qy 2821 AATGAAATTTACGACATCATCAACAGAAATTTAGAGTGTCTGTCTTCATCTTTG 2880  
Db 3108 AATGAAATTTACGACATCATCAACAGAAATTTAGAGTGTCTGTCTTCATCTTTG 2880  
Qy 2881 AATGTTCTATCCGACTCTCTTAATTTTCATTGA 2913  
Db 3168 AATGTTCTATCCGACTCTCTTAATTTTCATTGA 3200

## RESULT 5

US-10-425-114-26254  
; Sequence 26254, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313) B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 26254  
; LENGTH: 2836  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4119-112-B4\_FLI  
US-10-425-114-26254

Query Match 95.7%; Score 2786.4; DB 13; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2787; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 126 GATGATTAAGAAAGCCATGTACAAAGCAGAAATGGTACAGAGATCCAAATAGAGTGA 185  
Db 1 GATGATTAAGAAAGCCATGTACAAAGCAGAAATGGTACAGAGATCCAAATAGAGTGA 60  
Qy 186 AATGATTCGCAATGGAACATCCTCTTATCTTGGAGCTTTTAACTATTTGAAGATG 245  
Db 61 AATGATTCGCAATGGAACATCCTCTTATCTTGGAGCTTTTAACTATTTGAAGATG 120  
Qy 246 CAATTATGTATCTGTATTAAGAAATGTGCATATGAGAAATGAAAGAGTATCTAA 305  
Db 121 CAATTATGTATCTGTATTAAGAAATGTGCATATGAGAAATGAAAGAGTATCTAA 180  
Qy 306 GAATGAGTGAACCCCTTCTCAGAAATGAAGCTGACATTCATGACCAAGATCATC 365  
Db 181 GAATGAGTGAACCCCTTCTCAGAAATGAAGCTGACATTCATGACCAAGATCATC 240  
Qy 366 AGGATGTTGTATCTTCAATCTCATGTATCTACACGGGAGCTTCAACATTTCTAACCT 425  
Db 241 AGGATGTTGTATCTTCAATCTCATGTATCTACACGGGAGCTTCAACATTTCTAACCT 300  
Qy 426 CTTACTGACTCGTAAATATGAACATCAAGATTGCTGATTTGGGCTGGGAACTCAACTGAA 485  
Db 301 CTTACTGACTCGTAAATATGAACATCAAGATTGCTGATTTGGGCTGGGAACTCAACTGAA 360  
Qy 486 AATGCAATGAAAGCACTATACCTTATGAGAACTCTTAATCACTTCAACAGAAAT 545  
Db 361 AATGCAATGAAAGCACTATACCTTATGAGAACTCTTAATCACTTCAACAGAAAT 420

Qy 546 TGGCACTCGAAGTGCACATGGCCCTGAATCTGATGTTTGGTCCCTGGGCTGTATGTTTTA 605  
Db 421 TGGCACTCGAAGTGCACATGGCCCTGAATCTGATGTTTGGTCCCTGGGCTGTATGTTTTA 480  
Qy 606 TACATTACTTATCGGAGACCACTTGCAGACTGACACAGTCAGAAACATTAATTA 665  
Db 481 TACATTACTTATCGGAGACCACTTGCAGACTGACACAGTCAGAAACATTAATTA 540  
Qy 666 AGTAGATTTGGAGATTAATGAATGCCATCTTTTGTCAATAGAGCCAGACCTTAT 725  
Db 541 AGTAGATTTGGAGATTAATGAATGCCATCTTTTGTCAATAGAGCCAGACCTTAT 600  
Qy 726 TCACCAAGTACTGATGAGAAATCCAGACATGCTTAAGCTGTCTTCAATTTGAGCA 785  
Db 601 TCACCAAGTACTGATGAGAAATCCAGACATGCTTAAGCTGTCTTCAATTTGAGCA 660  
Qy 786 TCCCTTTATGTCGCCAAATTCCTTCAACAAAAGTAAAGATTAGAACTGTGAGACTC 845  
Db 661 TCCCTTTATGTCGCCAAATTCCTTCAACAAAAGTAAAGATTAGAACTGTGAGACTC 720  
Qy 846 AATGATATGAGGACATGCGACAAATTTCTAGCAATTAACAGCTTCTCAAGTACAGTAT 905  
Db 721 AATGATATGAGGACATGCGACAAATTTCTAGCAATTAACAGCTTCTCAAGTACAGTAT 780  
Qy 906 AAGTGTAGTTTATTTGACAAAGAACCTTTGATGGTAGGCACTCCCAATTAAT 965  
Db 781 AAGTGTAGTTTATTTGACAAAGAACCTTTGATGGTAGGCACTCCCAATTAAT 840  
Qy 966 GACTGATATTTCCAAAGATTAAGTTCACTGATTTTCTTCCAGAGATGAGAAACAG 1025  
Db 841 GACTGATATTTCCAAAGATTAAGTTCACTGATTTTCTTCCAGAGATGAGAAACAG 900  
Qy 1026 TTTTATCTCAGTGGGGAATTCAGAAACCAAGTAATGTGAGAGGGGAGAGTAATCA 1085  
Db 901 TTTTATCTCAGTGGGGAATTCAGAAACCAAGTAATGTGAGAGGGGAGAGTAATCA 960  
Qy 1086 AGATGCAAGAAAGGCCACATCTCGATACCTTCGTAGAGCTTATTCCTGTAGATC 1145  
Db 961 AGATGCAAGAAAGGCCACATCTCGATACCTTCGTAGAGCTTATTCCTGTAGATC 1020  
Qy 1146 TGGCACTTATATGTGATGCTCAAGCAAAACATATTAATGAGATGATCTCAGC 1205  
Db 1021 TGGCACTTATATGTGATGCTCAGCAAAACATATTAATGAGATGATCTCAGC 1080  
Qy 1206 AGAATGCTTCACTGATCCAAAGATCAGGAGGAGGAAATGAAAGAGTACTCACC 1265  
Db 1081 AGAATGCTTCACTGATCCAAAGATCAGGAGGAGGAAATGAAAGAGTACTCACC 1140  
Qy 1266 CACAGACAACATGCCAACAATTTTAACTTCTTAAAGAAAGACATCCAGTATCTGG 1325  
Db 1141 CACAGACAACATGCCAACAATTTTAACTTCTTAAAGAAAGACATCCAGTATCTGG 1200  
Qy 1326 ATCTTTGAAAGACCTGATTAACATCAAGACTCTCAATCATCTTTGTCCAGAAAAAC 1385  
Db 1201 ATCTTTGAAAGACCTGATTAACATCAAGACTCTCAATCATCTTTGTCCAGAAAAAC 1260  
Qy 1386 TCCCTTTCACTTGGCAGACCCGACACTCAGACTGAAACCGTCAACAGGTTTGGGAA 1445  
Db 1261 TCCCTTTCACTTGGCAGACCCGACACTCAGACTGAAACCGTCAACAGGTTTGGGAA 1320  
Qy 1446 TCTGCAAAATATATGCTCATTTAAGAAAAAAGTACTGATATGACAGATCAGCCAAACCG 1505  
Db 1321 TCTGCAAAATATATGCTCATTTAAGAAAAAAGTACTGATATGACAGATCAGCCAAACCG 1380  
Qy 1506 GGACTTCAGGGGCAATCCAGATTTGAGAAAGCAATCAAAAAAATGCTTGAATGAC 1565  
Db 1381 GGACTTCAGGGGCAATCCAGATTTGAGAAAGCAATCAAAAAAATGCTTGAATGAC 1440  
Qy 1566 AAAAGTCAAAAAGACCTGATGCTTCTGATTAATGACATCTGTAAACAGCAAAATGAC 1625  
Db 1441 AAAAGTCAAAAAGACCTGATGCTTCTGATTAATGACATCTGTAAACAGCAAAATGAC 1500  
Qy 1626 CATGAAATATATGACTGCACTTCAAGTAAACCTGATATTCACAAAGAAATGTTT 1685

Db 1501 CATGAATATATATACGACCTTACAGTAACCTGAGATATCAACAAGAAATGTGTTT 1560  
 Qy 1686 TGGCTCAGATCCTCTTCTTGAACAGCAAGACTAGGGGTAATGAGCCACCATGGGTTA 1745  
 Db 1561 TGGCTCAGATCCTCTTCTTGAACAGCAAGACTAGGGGTAATGAGCCACCATGGGTTA 1620  
 Qy 1746 TGAAGATGATACATTAAGAGCAATCACTCCGTTGTTGCTCAGAGGTTAAACCAAT 1805  
 Db 1621 TGAAGATGATACATTAAGAGCAATCACTCCGTTGTTGCTCAGAGGTTAAACCAAT 1680  
 Qy 1806 CAGACAGAAAACCAAAAAGGCTGTGTGTGACATCTTGAATTCAGAGAGGTGTGTGA 1865  
 Db 1681 CAGACAGAAAACCAAAAAGGCTGTGTGTGACATCTTGAATTCAGAGAGGTGTGTGA 1740  
 Qy 1866 GCTTGAAGAGATGATCTCAAGAAATGTGAAGAAGTTCTTCAGATATCTAGTA 1925  
 Db 1741 GCTTGAAGAGATGATCTCAAGAAATGTGAAGAAGTTCTTCAGATATCTAGTA 1800  
 Qy 1926 TGAAGATGATACATTAATTAATGATGAGGTTTCCCTTGTGATAG 1985  
 Db 1801 TGAAGATGATACATTAATTAATGATGAGGTTTCCCTTGTGATAG 1860  
 Qy 1986 ACCACCTCAGCTACTGACAAACATCAGTAGTACAGCTTGAACAATTTACAGAAAAA 2045  
 Db 1861 ACCACCTCAGCTACTGACAAACATCAGTAGTACAGCTTGAACAATTTACAGAAAAA 1920  
 Qy 2046 CTGGGCAAAATATCAATATGCTCCAGGTTTGTACAGCTTTGAAGTCTAAATCTCCAA 2105  
 Db 1921 CTGGGCAAAATATCAATATGCTCCAGGTTTGTACAGCTTTGAAGTCTAAATCTCCAA 1980  
 Qy 2106 AATCACTTATTTTACAAGATGCTAAATGCTAATTTGATGAGAAATTCCTGTGTGCTGA 2165  
 Db 1981 AATCACTTATTTTACAAGATGCTAAATGCTAATTTGATGAGAAATTCCTGTGTGCTGA 2040  
 Qy 2166 TTTTGAAGTTGTTTATGATGAGGTTAAATAATACAAACAAGAAATTTCAAGGT 2225  
 Db 2041 TTTTGAAGTTGTTTATGATGAGGTTAAATAATACAAACAAGAAATTTCAAGGT 2100  
 Qy 2226 GATTGAAGAGACAGGAAAGTCTTACCTTTAAAGAGTAAGTAAATGCTTGA 2285  
 Db 2101 GATTGAAGAGACAGGAAAGTCTTACCTTTAAAGAGTAAGTAAATGCTTGA 2160  
 Qy 2286 AAGAGATTAATAATGATATGAGCAAGCTAATGAGGTCATGATTTGTTAGCACT 2345  
 Db 2161 AAGAGATTAATAATGATATGAGCAAGCTAATGAGGTCATGATTTGTTAGCACT 2220  
 Qy 2346 GGAATTCATATTTTCAAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCAATAT 2405  
 Db 2221 GGAATTCATATTTTCAAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCAATAT 2280  
 Qy 2406 CATAGAGAGAAAACTGTGATGATCTAAGCTTACCTAAGCCTTATCACTCTCTCTGT 2465  
 Db 2281 CATAGAGAGAAAACTGTGATGATCTAAGCTTACCTAAGCCTTATCACTCTCTCTGT 2340  
 Qy 2466 GGAATTCATATTTTCAAGAGAGAAAGTAAGCAATTTCAACAAGATGCTATGCTG 2525  
 Db 2341 GGAATTCATATTTTCAAGAGAGAAAGTAAGCAATTTCAACAAGATGCTATGCTG 2400  
 Qy 2526 TGCCTTCTCAACAGAGCAACCAATCTTATCCCTTATGTTACAAATGAAGAGATTGG 2585  
 Db 2401 TGCCTTCTCAACAGAGCAACCAATCTTATCCCTTATGTTACAAATGAAGAGATTGG 2460  
 Qy 2586 TCTTACACTACAGCTTGTGAGAAAGACATCTTCTAATAGTCTAAGAAATGTCTTCC 2645  
 Db 2461 TCTTACACTACAGCTTGTGAGAAAGACATCTTCTAATAGTCTAAGAAATGTCTTCC 2520  
 Qy 2646 TAAATACAGCAAACTTTTGAATCTGTTTGTGAAGAAATGTTGGTGTGCAACAGTT 2705  
 Db 2521 TAAATACAGCAAACTTTTGAATCTGTTTGTGAAGAAATGTTGGTGTGCAACAGTT 2580  
 Qy 2706 AACTAGTGAAGCTGTGTGAGTTCAAGTTAATGATGAGGTCAGAGTTGGTGTGCAAGCAG 2765

Db 2581 AACTAGTGAAGCTGTGTGAGTTCAAGTTAATGATGAGGTCAGAGTTGTGCAAGCAG 2640  
 Qy 2766 AGTGTCTTATCAGTTATATCTACCAAAATGCTCAAAACATAGGTATGAGAAATGA 2825  
 Db 2641 AGTGTCTTATCAGTTATATCTACCAAAATGCTCAAAACATAGGTATGAGAAATGA 2700  
 Qy 2826 AAAATTACAGACTACATCAACAGAAATTAAGTGTCTGCTTCCATCTTTGATGTT 2885  
 Db 2701 AAAATTACAGACTACATCAACAGAAATTAAGTGTCTGCTTCCATCTTTGATGTT 2760  
 Qy 2886 TTCTAATCCGACTCCTAATTTTCAATGA 2913  
 Db 2761 TTCTAATCCGACTCCTAATTTTCAATGA 2788

RESULT 6  
 US-10-062-674-1868  
 ; Sequence 1868, Application US/10062674  
 ; Publication No. US2004000559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loring, Jeanne P.; Kager, Matthew R.  
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 ; FILE REFERENCE: PA-0026-1 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/062,674  
 ; PRIOR FILING DATE: 2002-01-30  
 ; PRIOR APPLICATION NUMBER: US 09/625,102  
 ; NUMBER OF SEQ ID NOS: 2217  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 1868  
 ; LENGTH: 3924  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US2004000559A1 332518.2  
 US-10-062-674-1868

Query Match 92.2%; Score 2686.8; DB 16; Length 3924;  
 Best Local Similarity 98.3%; Pred. No. 0;  
 Matches 2886; Conservative 0; Mismatches 27; Indels 23; Gaps 16;

Qy 1 ATGGGCACTGTCAGGGGAGAAATCGAGATTTTAAAGTTGAATCTGCTGTTAA 60  
 Db 285 ATGGGCACTGTCAGGGGAGAAATCGAGATTTTAAAGTTGAATCTGCTGTTAA 348  
 Qy 61 GGATCATTTGCTGTGTCTACAGAGCTGATCATTCACTGTGTTGGAAGTTGCAAT- 119  
 Db 349 GGATCATTTGCTGTGTCTACAGAGCTGATCATTCACTGTGTTGGAAGTTGCAAT- 408  
 Qy 120 CAAATGATATGATTAAGAAAGCCATGTACAAAGCAGAAATGTATCAGAGTCCAAATGA 179  
 Db 409 CAAATGATATGATTAAGAAAGCCATGTATCAAAAGCAGAAATGTATCAGAGTCCAAATGA 468  
 Qy 180 GGGAAATATACATGTCATTTGAAACATCTTCTATCTTGGAGCTTTATATATTTGA 239  
 Db 469 GGGAAATATACATGTCATTTGAAACATCTTCTATCTTGGAGCTTTATATATTTGA 528  
 Qy 240 AGATAGCAATTTATGTTATCTGTATTAAGAAATGTGCCATTAATGAGAAATGAACAGTA 299  
 Db 529 AGATAGCAATTTATGTTATCTGTATTAAGAAATGTGCCATTAATGAGAAATGAACAGTA 588  
 Qy 300 TCTTAAAGATATAGTGAAGAAACCTTCTCAGAAATTAAGTGTGACACTTATGACCAAGT 359  
 Db 589 TCTTAAAGATATAGTGAAGAAACCTTCTCAGAAATTAAGTGTGACACTTATGACCAAGT 648  
 Qy 360 CATCAAGGAGTGTATCTTCAATTCATGATGATATCAACCGGAGCTCACAATTTC 419  
 Db 649 CATCAAGGAGTGTATCTTCAATTCATGATGATATCAACCGGAGCTCACAATTTC 708  
 Qy 420 TAACTCTTACTGACTCGTAAATATGAACATCAAGATGTGCTGATTTTGGCTGGCACTCA 479  
 Db 709 TAACTCTTACTGACTCGTAAATATGAACATCAAGATGTGCTGATTTTGGCTGGCACTCA 768

QY 480 ACTGAAAATGCGACATGAAAAGCACTATACATTATGTGGAATCTCTTAACATTCATTTCACC 539  
 Db 769 ACTGAAAATGCGACATGAAAAGCACTATACATTATGTGGAATCTCTTAACATTCATTTCACC 828  
 QY 540 AGAATATGCGACTCGAAGTGCACATGCGCTTGAATCTGATGTGGTCCCTGGGCTGTAT 599  
 Db 829 AAAAAATGCGACTCGAAGTGCACATGCGCTTGAATCTGATGTGGTCCCTGGGCTGTAT 888  
 QY 600 GTTTATATACATTACTATTCGGGAGACCACTCTTGCACCTGACACATGCAAGCAACATTC 659  
 Db 889 GTTTATATACATTACTATTCGGGAGACCACTCTTGCACCTGACACATGCAAGCAACATTC 948  
 QY 660 AATAAATGATGATTGGCAGATTATGAAATGCACTCTTTTGTCAATATGAGGGCAAGGA 719  
 Db 949 AATAAATGATGATTGGCAGATTATGAAATGCACTCTTTTGTCAATATGAGGGCAAGGA 1008  
 QY 720 CCTTATTCACAGATTACTCTGTAAGAAATCCAGCAATGCTTTAAAGTCTGCTTCAGTATTC 779  
 Db 1009 CCTTATTCACAGATTACTCTGTAAGAAATCCAGCAATGCTTTAAAGTCTGCTTCAGTATTC 1068  
 QY 780 GGACCATCTCTTTTATGTCCGAAATCTTCACAAAAGTAAAGATTAGAACTGTGA 839  
 Db 1069 GGACCATCTCTTTTATGTCCGAAATCTTCACAAAAGTAAAGATTAGAACTGTGA 1128  
 QY 840 AGACTCAATTGATGATGGGCGATGCGCAATTTCTACTGCAATTACAGCTTCTTCAGTATC 899  
 Db 1129 AGACTCAATTGATGATGGGCGATGCGCAATTTCTACTGCAATTACAGCTTCTTCAGTATC 1188  
 QY 900 CAGTATAGTGTAGTATTATTTTGAAGAAAAGCTTTTGAATGTGTCAGCACTCCCAA 959  
 Db 1189 CAGTATAGTGTAGTATTATTTTGAAGAAAAGCTTTTGAATGTGTCAGCACTCCCAA 1248  
 QY 960 TAAAAATGCTGTATTTCCAAAAGATAAAGTCACTGATTTTCTTCTTCAGAGATGG 1019  
 Db 1249 TAAAAATGCTGTATTTCCAAAAGATAAAGTCACTGATTTTCTTCTTCAGAGATGG 1308  
 QY 1020 AAACAGTTTATATCAAGTGGGGAATCAAGAAACCATATATGTGGAAGGGGAAGT 1079  
 Db 1309 AAACAGTTTATATCAAGTGGGGAATCAAGAAACCATATATGTGGAAGGGGAAGT 1368  
 QY 1080 AATTCAAGATGCAAGAAAAGGCCCACTTCGATACCTTCGATAGCTTATTCCTGGA 1139  
 Db 1369 AATTCAAGATGCAAGAAAAGGCCCACTTCGATACCTTCGATAGCTTATTCCTGGA 1428  
 QY 1140 TAGATCTGGCACTTCAATATGTCAGTCTCAAGCAAAAACATATCAATGGAACGATGCA 1199  
 Db 1429 TAGATCTGGCACTTCAATATGTCAGTCTCAAGCAAAAACATATCAATGGAACGATGCA 1488  
 QY 1200 CTCAGCAAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAAAATGGAAGAGGTA 1259  
 Db 1489 CTCAGCAAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAAAATGGAAGAGGTA 1548  
 QY 1260 CTCACCCACAGACAAATGCGCAACTTTTAACTTTTAAAGAAAAGACATCCAGTAG 1319  
 Db 1549 CTCACCCACAGACAAATGCGCAACTTTTAACTTTTAAAGAAAAGACATCCAGTAG 1608  
 QY 1320 TTCTGATCTTTTGAAGAGCTGATTAACAATCAAGCACTCCCAAT-CATCTTGTCCAG 1378  
 Db 1609 TTCTGATCTTTTGAAGAGCTGATTAACAATCAAGCACTCCCAAT-CATCTTGTCCAG 1668  
 QY 1379 GAAAAATCTCTTTTCAATTTGAGACCCGAGCACTCAGACTGAAACGTTAACAAGTGT 1438  
 Db 1669 GAAAAATCTCTTTTCAATTTGAGACCCGAGCACTCAGACTGAAACGTTAACAAGTGT 1728  
 QY 1439 TTGGGAATCTGCAATAATGCTCATTTAAGAAAACATCTGAAATATGACACGATCC 1498  
 Db 1729 TTGGGAATCTGCAATAATGCTCATTTAAGAAAACATCTGAAATATGACACGATCC 1788  
 QY 1499 CAAACCGGGAATTCAGGGGCATCCAGATTTGCAAGAAAGCAATCAAAAAATGCCGGA 1558  
 Db 1789 CAAACCGGGAATTCAGGGGCATCCAGATTTGCAAGAAAGCAATCAAAAAATGCCGGA 1848

QY 1559 CTGATACAAAAGTCAAAAGAACTGTGATCTTGTATATGACATTTCTGTAACACG 1618  
 Db 1849 CTGATACAAAAGTCAAAAGAACTGTGATCTTGTATATGACATTTCTGTAACACG 1908  
 QY 1619 AAAAAATCAATGAATATATGACTGCACTTCACAGTAAACTGAGATTAATCCAAAGAAAT 1678  
 Db 1909 AAAAAATCAATGAATATATGACTGCACTTCACAGTAAACTGAGATTAATCCAAAGAAAT 1968  
 QY 1679 GTGTGTTGGCTCAGATCCTCTTTCTGAAACAGAGCAACTGAGGGATATGAGCCACAT 1738  
 Db 1969 GTGTGTTGGCTCAGATCCTCTTTCTGAAACAGAGCAACTGAGGGATATGAGCCACAT 2028  
 QY 1739 GGGGTATCAGATCGTATCATTTAGAGCAATTAACATCCCTGGTGTCTCAAGGTTAA 1798  
 Db 2029 GGGGTATCAGATCGTATCATTTAGAGCAATTAACATCCCTGGTGTCTCAAGGTTAA 2088  
 QY 1799 AACCAATCAGACAGAAAACCAAAAAGCTGTGGTGAAGATACCTTGATTCAGAGAGGTGT 1858  
 Db 2089 AACCAATCAGACAGAAAACCAAAAAGCTGTGGTGAAGATACCTTGATTCAGAGAGGTGT 2148  
 QY 1859 GTGTGAAGCTGTAAAGAGATGATCTCAAGAATATGTGAAAGAGTCTTCAGATAT 1918  
 Db 2149 GTGTGAAGCTGTAAAGAGATGATCTCAAGAATATGTGAAAGAGTCTTCAGATAT 2208  
 QY 1919 CTAGATGAGAAAATACATCACTATTTATTCCAAATGCT-GGTAGAGTTCCTCTT 1977  
 Db 2209 CTAGATGAGAAAATACATCACTATTTATTCCAAATGCTGAGTTCCTCTT 2268  
 QY 1978 GCT-GATAGACCACTTCACCTAC-TGACAAACATCAATAGGTACAGTTGACATTTAC 2035  
 Db 2269 GCTGATGAGCAACCTTCACCTACCTTACCAACATCAATAGGTACAGTTGACATTTAC 2328  
 QY 2036 CAGAAAATATCTGGCGAAAATATCATATATGCTTCAGGTTGTACAGCTGTAAAGATCTA 2095  
 Db 2329 CAGAAAATATCTGGCGAAAATATCATATATGCTTCAGGTTGTACAGCTGTAAAGATCTA 2388  
 QY 2096 AATCTCCCAAATACATTAATTTTACAAAGATATGCAATTTTGAATGAGAAATCTC 2155  
 Db 2389 AATCTCCCAAATACATTAATTTTACAAAGATATGCAATTTTGAATGAGAAATCTC 2448  
 QY 2156 CTGGTGTGATTTTGAAGTGTGTTTATGATGGGTAAAAAATACAAAACAGAGATT 2215  
 Db 2449 CTGGTGTGATTTTGAAGTGTGTTTATGATGGGTAAAAAATACAAAACAGAGATT 2508  
 QY 2216 TCATTCAGAGATTTGAAAAGACAGGGAAGTCTTACCTTTAAAAAAGTGAAGTTA 2275  
 Db 2509 TCATTCAGAGATTTGAAAAGACAGGGAAGTCTTACCTTTAAAAAAGTGAAGTTA 2568  
 QY 2276 ATAGCTTGAAGAAGATTAATAATGTAT-ATGACCATGCTAAT-GAGGTCATCGTAT 2333  
 Db 2569 ATAGCTTGAAGAAGATTAATAATGTATATGACCATGCTAATGAGGTCATCGTAT 2628  
 QY 2334 TTGTATAGACCTGGAAATCCATATTTTCAAGAGAGAAAAGAAAACCTAGAGTGTCCCTT 2393  
 Db 2629 TTGTATAGACCTGGAAATCCATATTTTCAAGAGAGAAAAGAAAACCTAGAGTGTCCCTT 2688  
 QY 2394 TTTCCTCAATATCATAGAGAAAACCTGGT-AGTACTAGTTCACTTAAGGCTTATAC 2452  
 Db 2689 TTTCCTCAATATCATAGAGAAAACCTGGT-AGTACTAGTTCACTTAAGGCTTATAC 2748  
 QY 2453 CTCCTCTCTT-TCGTGATTCAAATTAACCAACGAGAGATAGACATCTTTCAACAGATG 2511  
 Db 2749 CTCCTCTCTT-TCGTGATTCAAATTAACCAACGAGAGATAGACATCTTTCAACAGATG 2808  
 QY 2512 GTCAATGATAGTGTGCTTCTCCAAACAGAGCAACAACTCTTAATCCCTCTATGTTACA 2571  
 Db 2809 GTCAATGATAGTGTGCTTCTCCAAACAGAGCAACAACTCTTAATCCCTCTATGTTACA 2868  
 QY 2572 AATGAAGACTTGTGCTTAAACAATGCTCT-GGAACAGACATCTCTTAATGCT 2630  
 Db 2869 AATGAAGACTTGTGCTTAAACAATGCTCTGGAACAGACATCTCTTAATGCT 2928  
 QY 2631 AAAAGATGTCTTCTTA---TCAGCAACAATTTGAATC--TGTGTTTGTGAAAAA 2684

|||||  
Db 2929 AAAAGATGTCCTTCCATATATACGACCAACCTTTGAAATCGTTTTGTGAAAAAAT 2988  
|||  
Qy 2685 TGTGTGGTGGCTAC-----CAGTTAACTAGTGGAGCTGTGTGGCTTCACTTTAATGATG 2740  
|||  
Db 2989 GTTTGGTTGGGGCATACCAACAGTTAACTAGTGGAGCTGTGTGGCTTCACTTTAATGATG 3048  
|||  
Qy 2741 GGTCCAGTTGGTGTGGAGGGAG--GAGTGTCTTATC-AGTTATACCTCACCAATGG 2798  
|||  
Db 3049 GGTCCAGTTGGTGTGGAGGGAGAGTGTCTTATACAGTTATACCTCACCAATGG 3108  
|||  
Qy 2799 TCAA-CAACTAGATGAGGAAATGAAATTAACAGACTACATCAACAGAAATTAAC 2857  
|||  
Db 3109 TCAACCAACTAGATGAGGAAATGAAATTAACAGACTACATCAACAGAAATTAAC 3168  
|||  
Qy 2858 AGTGTGTCTTCCATCCCTTTTGAATGTTTCTAATCGACTCTTAATTTTCAATGA 2913  
|||  
Db 3169 AGTGTGTCTTCCATCCCTTTTGAATGTTTCTAATCGACTCTTAATTTTCAATGA 3224  
|||

## RESULT 7

US-09-918-995-21560  
; Sequence 21560, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hybreg, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21560  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(484)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-21560

Query Match 14.6%; Score 426; DB 10; Length 484;  
Best Local Similarity 100.0%; Pred. No. 5.2e-101;  
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2026 GACAAATTTACGAAATATACGCGGAAATATCAATATGCTCCAGGTTTGTACAGCTT 2085  
|||  
Db 59 GACAAATTTACGAAATATACGCGGAAATATCAATATGCTCCAGGTTTGTACAGCTT 118  
|||  
Qy 2086 GTAAAGATCTAAATCTCCAAATCACTTAATTTTACAAGATATGCTAAATGATTTGATG 2145  
|||  
Db 119 GTAAAGATCTAAATCTCCAAATCACTTAATTTTACAAGATATGCTAAATGATTTGATG 178  
|||  
Qy 2146 GAGAATTTCTCGTGCTGCTGATTTTGAAGTTTGTATGATGGGTAAATATACAAA 2205  
|||  
Db 179 GAGAATTTCTCGTGCTGCTGATTTTGAAGTTTGTATGATGGGTAAATATACAAA 238  
|||  
Qy 2206 ACAGAAATTTCACTTACAGATGATGAAAGACAGGAAAGCTTCACTTTAAATATGAA 2265  
|||  
Db 239 ACAGAAATTTCACTTACAGATGATGAAAGACAGGAAAGCTTCACTTTAAATATGAA 298  
|||  
Qy 2266 AGTGAATTAATAGCTTGAAGAGAGATTAATTAATGATGAGCATGCTAATGAGGT 2325  
|||  
Db 299 AGTGAATTAATAGCTTGAAGAGAGATTAATTAATGATGAGCATGCTAATGAGGT 358  
|||  
Qy 2326 CATCGAATTTGTTTGAAGCTGGAATCCATTAATTTCAAGAGAGAAAGAAATCAGAGT 2385  
|||  
Db 359 CATCGAATTTGTTTGAAGCTGGAATCCATTAATTTCAAGAGAGAAAGAAATCAGAGT 418  
|||

Qy 2386 GCTCCCTTTTCCCAATATCATAGAGAAAACTGTAGTACTAGTTCACCTAAGGCC 2445  
|||  
Db 419 GCTCCCTTTTCCCAATATCATAGAGAAAACTGTAGTACTAGTTCACCTAAGGCC 478  
|||  
Qy 2446 TTATCA 2451  
|||  
Db 479 TTATCA 484  
|||

## RESULT 8

US-09-878-178-96  
; Sequence 96, Application US/09878178  
; Patent No. US2002017752A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527  
; CURRENT APPLICATION NUMBER: US/09/878,178  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 2237  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 96  
; LENGTH: 328  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-878-178-96

Query Match 11.3%; Score 328; DB 9; Length 328;  
Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1428 ACAACAGTGTGTTGGAAATCTGCAATTAATATCTCATTTAAGAAAACTAGTAATATGA 1487  
|||  
Db 1 ACAACAGTGTGTTGGAAATCTGCAATTAATATCTCATTTAAGAAAACTAGTAATATGA 60  
|||  
Qy 1488 CAGCATCAGCCCAACCGGAGCTTCCAGGCGCATCCAGATTGGCAAGACATCAAA 1547  
|||  
Db 61 CAGCATCAGCCCAACCGGAGCTTCCAGGCGCATCCAGATTGGCAAGACATCAAA 120  
|||  
Qy 1548 AATGCTGAGTGTATCAAAAGTCAAAAAGACTGTGATGCTTGTATATGACATTC 1607  
|||  
Db 121 AATGCTGAGTGTATCAAAAGTCAAAAAGACTGTGATGCTTGTATATGACATTC 180  
|||  
Qy 1608 TGTAAACAGCAAAATATACATGAAATATATGATGATCTTCAAGTAACTGAGATTAAT 1667  
|||  
Db 181 TGTAAACAGCAAAATATACATGAAATATATGATGATCTTCAAGTAACTGAGATTAAT 240  
|||  
Qy 1668 CCAACAAAGATGTTTGTGCTCAGATCTCTTTCTGAACAGAGAAAGCTAGGGGTAT 1727  
|||  
Db 241 CCAACAAAGATGTTTGTGCTCAGATCTCTTTCTGAACAGAGAAAGCTAGGGGTAT 300  
|||  
Qy 1728 GGAGCCACATGGGTTATCAGATCGT 1755  
|||  
Db 301 GGAGCCACATGGGTTATCAGATCGT 328  
|||

## RESULT 9

US-10-046-935-96  
; Sequence 96, Application US/10046935  
; Publication No. US20020156011A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527C1  
; CURRENT APPLICATION NUMBER: US/10/046,935

/ CURRENT FILING DATE: 2002-01-15  
/ NUMBER OF SEQ ID NOS: 2239  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 96  
/ LENGTH: 328  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-10-046-935-96

Query Match 11.3%; Score 328; DB 14; Length 328;  
Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 ACAACAGTGTGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 1487  
DB 1 ACAACAGTGTGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 60  
QY 1488 CAGCATCAGCCCAAAACCGGCACTTCAGGGCCATTCAGATTGTCAGAAAGACATCAAA 1547  
DB 61 CAGCATCAGCCCAAAACCGGCACTTCAGGGCCATTCAGATTGTCAGAAAGACATCAAA 120  
QY 1548 AATGCTGAGCTATACAAAAGTCAAAAAGACTGAGCTTCTGATATGCACTTC 1607  
DB 121 AATGCTGAGCTATACAAAAGTCAAAAAGACTGAGCTTCTGATATGCACTTC 180  
QY 1608 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAGAGTAACTGAGATAT 1667  
DB 181 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAGAGTAACTGAGATAT 240  
QY 1668 CCAACAAGATGTGTTTGGCTCAGATCCTCTTCTTGAAACAGACGACATGAGGATAT 1727  
DB 241 CCAACAAGATGTGTTTGGCTCAGATCCTCTTCTTGAAACAGACGACATGAGGATAT 300  
QY 1728 GAGGCCACCATGGGGTTATCAGATCGT 1755  
DB 301 GAGGCCACCATGGGGTTATCAGATCGT 328

RESULT 10  
US-10-146-502-96  
/ Sequence 96, Application US/10146502  
/ Publication No. US20030069180A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Jiang, Yudi  
/ APPLICANT: Harlocker, Susan L.  
/ APPLICANT: Secrist, Heather  
/ APPLICANT: Wang, Aijun  
/ APPLICANT: Stolk, John A.  
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
/ FILE REFERENCE: 210121.527C2  
/ CURRENT APPLICATION NUMBER: US/10/146,502  
/ CURRENT FILING DATE: 2002-05-14  
/ NUMBER OF SEQ ID NOS: 2241  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 96  
/ LENGTH: 328  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-10-146-502-96

Query Match 11.3%; Score 328; DB 15; Length 328;  
Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 ACAACAGTGTGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 1487  
DB 1 ACAACAGTGTGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 60  
QY 1488 CAGCATCAGCCCAAAACCGGCACTTCAGGGCCATTCAGATTGTCAGAAAGACATCAAA 1547  
DB 61 CAGCATCAGCCCAAAACCGGCACTTCAGGGCCATTCAGATTGTCAGAAAGACATCAAA 120

QY 1548 AATGCTGAGCTATACAAAAGTCAAAAAGACTGAGCTTCTGATATGCACTTC 1607  
DB 121 AATGCTGAGCTATACAAAAGTCAAAAAGACTGAGCTTCTGATATGCACTTC 180  
QY 1608 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAGAGTAACTGAGATAT 1667  
DB 181 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAGAGTAACTGAGATAT 240  
QY 1668 CCAACAAGATGTGTTTGGCTCAGATCCTCTTCTTGAAACAGACGACATGAGGATAT 1727  
DB 241 CCAACAAGATGTGTTTGGCTCAGATCCTCTTCTTGAAACAGACGACATGAGGATAT 300  
QY 1728 GAGGCCACCATGGGGTTATCAGATCGT 1755  
DB 301 GAGGCCACCATGGGGTTATCAGATCGT 328

RESULT 11  
US-10-062-674-914  
/ Sequence 914, Application US/10062674  
/ Publication No. US20040005559A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
/ TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
/ FILE REFERENCE: PA-0026-1 CIP  
/ CURRENT APPLICATION NUMBER: US/10/062,674  
/ CURRENT FILING DATE: 2002-01-30  
/ PRIOR APPLICATION NUMBER: US 09/625,102  
/ NUMBER OF SEQ ID NOS: 2217  
/ SOFTWARE: PERL Program  
/ SEQ ID NO 914  
/ LENGTH: 445  
/ TYPE: DNA  
/ ORGANISM: Mus musculus  
/ FEATURE:  
/ NAME/KEY: misc feature  
/ OTHER INFORMATION: GenBank ID No. US20040005559A1 g1919634  
US-10-062-674-914

Query Match 9.5%; Score 276; DB 16; Length 445;  
Best Local Similarity 81.2%; Pred. No. 1.3e-61;  
Matches 359; Conservative 0; Mismatches 75; Indels 8; Gaps 3;

QY 1576 AAGAATCTGATGCTCTGTAATATGACATCTCTGTAACAGCAAAATACATGAATAT 1635  
DB 12 AAGAATCTGATGCTCTGTAATATGACATCTCTGTAACAGCAAAATACATGAATAT 70  
QY 1636 ATGACTGCACTTACAGTAACTGAGATATATCAACAAGAAATGTTTGGCTCAGAT 1695  
DB 71 ATGAGTGCATCATCACCATTAAGCTGAGTCAATGCCACAGAA-----GCCGGCTCAAT 124  
QY 1696 CTTCTTCTGAAACAGCAAGACTAGGGTATGAGCCCAATGGGGTTATCAGATCGT 1755  
DB 125 CTTCTTCTGAAACAGCAAGACTAGGGTATGAGCCCAATGGGGTTATCAGATCGT 184  
QY 1756 ACATTAAGAAGCTTACATCTCGTGGTGTCTCAAGGTTAAACCAATGAGAGAAA 1815  
DB 185 ACCTTAAGAAGTATATACATCTCTGATGCTCAAGATTAAGCAATATCAGACAGAA 244  
QY 1816 ACCAAAAAGGCTGTGTGAGCATCTTGAATTCAGAGAGGTGTGTGAGCTTGAAG 1875  
DB 245 ACCAAAAAGGCTGTGTGAGCATCTTGAATTCAGAGAGGTGTGTGAGCTTGAAG 304  
QY 1876 GAGTATGATCTCAAGAAATATGTAAGAGTTCTTCAATATCTAGATGAGAAATGCG 1935  
DB 305 GAGTATGATCTCAAGAAATATGTAAGAGTTCTTCAATATCTAGATGAGAAATGCG 363  
QY 1936 ATGACTATTATATCAAAATGAGTGTGAGGTTTCTCTTGTGCTATGAGACCACTCA 1995  
DB 364 ATGACTATTATATCAAAATGAGTGTGAGGTTTCTCTTGTGCTATGAGACCTCTTG 423  
QY 1996 CCTACTGCAACATCAGTAGT 2017

Db 424 CCTACTGACGACATCTAGT 445

RESULT 12  
US-09-918-995-24717  
; Sequence 24717, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: HySeq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCES: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24717  
; LENGTH: 491  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(491)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-24717

Query Match 8.6%; Score 250; DB 10; Length 491;  
Best Local Similarity 100.0%; Pred. No. 9e-55;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGTTGTA 60  
DB 241 ATGGCGACCTGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGTTGTA 300  
QY 61 GGATCATTTGCTGCTGTCTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 120  
DB 301 GGATCATTTGCTGCTGTCTACAGAGCTGATCCATTCACATGCTTTGGAAGTTGCAATC 360  
QY 121 AAAATGATGATTAAGAAAGCATGTACAAAGCAGAAATGTTACAGAGTCCAAAATGAG 180  
DB 361 AAAATGATGATTAAGAAAGCATGTACAAAGCAGAAATGTTACAGAGTCCAAAATGAG 420  
QY 181 GTGAAATATCATTTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATTAACATTTTGA 240  
DB 421 GTGAAATATCATTTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATTAACATTTTGA 480  
QY 241 GATAGCAATT 250  
DB 481 GATAGCAATT 490

RESULT 13  
US-10-152-319A-292/c  
; Sequence 292, Application US/10152319A  
; Publication No. US20040072260A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Higgs, Brandon  
; APPLICANT: Castle, Arthur  
; APPLICANT: Blashoff, Michael  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCES: 44921-5089-US  
; CURRENT APPLICATION NUMBER: US/10/152,319A  
; PRIOR FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/292,335  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/297,523  
; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,925  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,810  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,807  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,808  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/315,047  
; PRIOR FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: US 60/324,928  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/330,867  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/330,462  
; PRIOR FILING DATE: 2001-10-22  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2221  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 292  
; LENGTH: 302  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. AA894059  
US-10-152-319A-292

Query Match 7.4%; Score 216.8; DB 12; Length 302;  
Best Local Similarity 91.3%; Pred. No. 3.4e-46;  
Matches 230; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 2662 TTGAATCTGTTTGTGAAATGTTGTTGGCTACACAGTTACTAGTGAGCTG 2721  
DB 302 TTGAATCTGTTTGTGAAATGTTGTTGGCTACACAGTTACTAGTGAGCTG 243  
QY 2722 TGGGTTCACTTAATGATGGTCCAGTTGTTGTGACAGGAGAGTCTTCTACAGT 2781  
DB 242 TGGGTTCACTTAATGATGGTCCAGTTGTTGTGACAGGAGAGTCTTCTACAGT 183  
QY 2782 TATACCTCACCAANTGCTAAACACTAGGATGAGAAATGAAAATTTACGACTAC 2841  
DB 182 TATACCTCACCAANTGCTAAACACTAGGATGAGAAATGAAAATTTACGACTAC 123  
QY 2842 ATCAACAGAAATTAAGGTGCTTCCATTCCTTTGATGTTTCTATCCGACTCCT 2901  
DB 122 ATCAACAGAAATTAAGGTGCTTCCATTCCTTTGATGTTTCTATCCGACTCCT 63  
QY 2902 AATTTTCATTGA 2913  
DB 62 AGTTTTCAGTAA 51

RESULT 14  
US-10-305-720-649  
; Sequence 649, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCES: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 649  
; LENGTH: 381  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040010136A1 259999

Search completed: September 24, 2004, 20:47:13  
Job time : 1312 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:57:40 ; Search time 15 Seconds

(without alignments)  
3367.203 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKEDPKVGNLGLK.....KIQCLSSILMLPSNPNNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599.5	11.8	576	1	POLO_DROME
2	593	11.7	631	1	CNK_MOUSE
3	590.5	11.6	615	1	060806 mus musculus
4	589.5	11.6	646	1	094044 homo sapien
5	561.5	11.1	682	1	SNK_MOUSE
6	560.5	11.0	682	1	SNK_MOUSE
7	560.5	11.0	685	1	SNK_MOUSE
8	554.5	10.9	407	1	STK6_HUMAN
9	548	10.8	408	1	STK6_XENLA
10	545	10.7	603	1	PLK1_MOUSE
11	537	10.6	603	1	PLK1_MOUSE
12	536	10.6	603	1	PLK1_MOUSE
13	526.5	10.4	1518	1	KKK1_YEAST
14	514.5	10.1	683	1	PLK1_MOUSE
15	508.5	10.0	403	1	STK6_HUMAN
16	504.5	9.9	632	1	PLK1_MOUSE
17	503.5	9.9	397	1	STK6_MOUSE
18	501.5	9.9	395	1	STK6_MOUSE
19	498	9.8	615	1	PLK1_MOUSE
20	498	9.8	661	1	ARK5_HUMAN
21	496.5	9.8	649	1	PLK1_MOUSE
22	496.5	9.8	649	1	PLK1_MOUSE
23	493.5	9.7	535	1	KI11_MOUSE
24	490	9.6	705	1	CDK5_YEAST
25	489.5	9.6	343	1	AURB_MOUSE
26	487.5	9.6	384	1	ARL1_MOUSE
27	486	9.6	1037	1	KCC4_YEAST
28	485.5	9.6	344	1	AURB_MOUSE
29	479.5	9.4	345	1	AURB_MOUSE
30	477.5	9.4	774	1	MRK2_MOUSE
31	471.5	9.3	282	1	AURC_MOUSE
32	469	9.2	367	1	IPK1_MOUSE
33	468	9.2	1081	1	KOKO_MOUSE

34	467.5	9.2	752	1	MRK4_HUMAN	Q96134 homo sapien
35	467	9.2	611	1	SNF1_CANCA	Q00372 candida gla
36	465.5	9.2	309	1	AURC_MOUSE	Q9ub9b homo sapien
37	462.5	9.1	633	1	SNF1_MOUSE	P06782 saccharomyc
38	454.5	9.0	556	1	PDPK_MOUSE	Q15530 homo sapien
39	454	8.9	559	1	PDPK_MOUSE	Q05173 rattus norv
40	453.5	8.9	620	1	SNF1_MOUSE	P52497 candida alb
41	450	8.9	559	1	PDPK_MOUSE	Q92240 mus musculu
42	449.5	8.9	794	1	KI11_MOUSE	Q81d3c homo sapien
43	447	8.8	776	1	MRK3_MOUSE	P27448 homo sapien
44	445	8.8	502	1	KI11_MOUSE	Q02723 secale cere
45	441.5	8.7	533	1	KCCD_MOUSE	P15791 rattus norv

## ALIGNMENTS

RESULT 1	ID	POLO_DROME	STANDARD	PRT	576 AA.
AC	P52304	Q9VWB2			
DT	01-OCT-1996	(Rel. 34, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Protein kinase polo (EC 2.7.1.-).				
GN	POLO OR CG12306.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Canton-S;				
RX	MEDLINE=92084090; PubMed=1660828;				
RA	Llamazares S., Moreira A., Tavares A., Girdham C., Spruce B.A.,				
RA	Gonzalez C., Karses R.E., Glover D.M., Sunkel C.E.;				
RT	"Polo encodes a protein kinase homolog required for mitosis in				
RT	Drosophila.";				
RL	Gene Dev. 5:2153-2165 (1991).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Berkley;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Maniatis P.G., Scher S.E., Li P.W., Hoskins R.A., Gale R.F.,				
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Branton R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,				
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abdel J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,				
RA	Bailly R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA	Burtis J.M., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.K., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Fosler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Glodok A., Goss F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,				
RA	Jalali M., Kalish F., Kapen G.H., Ke Z., Kempton J.A., Ketchum K.A.,				
RA	Kimble B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merklov G., Mleishina N.V., Moberg C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacht J.M.,				
RA	Palazzo M., Pittman G.S., Pan S., Pollard V., Puti V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,				

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
CC  
CC  
CC -I- FUNCTION: May play a role in regulating both nuclear and  
CC cytoplasmic aspects of the mitotic cycle.  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
CC  
CC -I- TISSUE SPECIFICITY: Larval discs, brain and testis.  
CC  
CC -I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC CDC5/Polo subfamily.  
CC  
CC -I- SIMILARITY: Contains 2 POLO box domains.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaborative  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; X63361; CAA44963.1; -.	
DR	EMBL; AE003514; AAF49036.1; --.	
DR	PIR; S22127; S22127.	
DR	HSSP; Q63450; 1A06.	
DR	FLYBase; FBgn0003124; polo.	
DR	GO; GO:0005813; C:centrosome; IDA.	
DR	GO; GO:0005819; C:spindle; IDA.	
DR	GO:0004674; P:protein serine/threonine kinase activity; IDA.	
DR	GO:0007714; P:male meiosis; IMP.	
DR	GO; GO:0007067; P:mitosis; IMP.	
DR	InterPro; IPR000959; POLO_box.	
DR	InterPro; IPR000719; Prot_kinase.	
DR	InterPro; IPR008271; Ser_thr_kin_AS.	
DR	InterPro; IPR002290; Ser_thr_pkinase.	
DR	Pfam; PF00069; pkinase; 1.	
DR	Pfam; PF00069; pkinase; 2.	
DR	ProDom; PD000001; Prot_kinase; 1.	
DR	SMART; SM00220; S_TKc; 1.	
DR	PROSITE; PS50078; POLO_BOX; 2.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.	
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
KW	Transferase; Serine/threonine protein kinase; ATP-binding; Repeat.	
FT	DOMAIN	25 .. 277
FT	NP_BINDING	31 .. 39
FT	BLINDING	54 .. 54
FT	ACT_SITE	148 .. 148
FT	DOMAIN	398 .. 461
FT	DOMAIN	496 .. 564
FT	CONFLICT	187 .. 187
SQ	SEQUENCE	576 AA; 56973 MW; 66973 MM; 5022B9AC0E888FAD CRC64; P -> A (IN REF. 1).

Query Match	11.8%	Score 539.5;	DB 1;	Length 576;
Best Local Similarity	34.1%;	Fred. No. 9.56-25;		
Matches 136;	Conservative 71;	Mismatches 121;	Indels 71;	Gaps 9;
QY	2	ATCIGEKIED-----FVGNILGKGSFAGYRABSHHTGLEVAIMIDKKMYTAGMGQ	55	
		::	::	:
Db	9	STDIDPRLVDINQRTKYGRMRPFQKSGSPACKYTEIIVDETIDVQVAGKIVSKRMITKNQKE	68	
QY	56	RQVNEVVKHCQKPSIELYNYFEDSNYYVLVLEMGCHENNRVLKNRFEFSENEARH	115	
		::	::	
Db	69	KTAQOEITTHRSINHNIVKPFNYFEDSQNIYIVLECKRSWMLKKRR-KSITFECKRY	127	
QY	116	FPHQIITGMALHSGILHRLDLSNLLTRNNAMIKADPGLATQAKMFHEKHYYLLCGTP	175	
		::	::	
Db	128	YIQLIQGVKYLHNDRIIHRDLKGNLFINDLLHAKVIGFGLATRIEYGEKRTKLTCTGP	187	
		::	::	

```

QY 176 MYISPELATSAGGLESDDWSLGCNEFYTLILGRPEPTDVTYKATLNVLIADYEMPSFLS 235
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 188 NYIAPELLITKKGHSFEVDIWSIGCWYITLVLVQGPPEFKTILKDYASKI KCEYRVPSYLR 247
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
QY 236 IEAKDLIHLRRNPDRLSLSSVLDHPMS-----RNSSTKSKDLG--TVYDST 283
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 248 KPAADWVIALQNPESRAPIQGLNFEFLKSKKYKMLPSSCLTMAPRGSDNTEDSM 307
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
QY 284 D-----SQ-----HATISTATASSTISGSLPDKRLLIGQLPKMT 323
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 308 HRPXLEMMNGIRPDRLESTFLKANLMDAITASQVC----- 345
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
QY 324 VFPKXKSTDPSSGGGSGNSFTYQMGQNSTNSGGRVYIQ 362
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 346 ----RHSEDDYRS--DISELYOQLTN--LINGPRILIQ 374
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

```

[illegible]

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL; U21392; AAC52191.1; -.  
DR EMBL; U22434; AAC52192.1; -.  
DR PIR; AF136586; AAF08369.1; -.  
DR PIR; A57286; A57286.  
DR MGD; MG1:109604; Cnk.  
DR InterPro; IPR000959; POLO\_box.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR008271; Ser\_thr\_kin\_AS.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00659; POLO\_box; 2.  
DR ProDom; PD000001; Prot\_Kinase; 1.  
DR SMART; SMO0220; S\_TKc; 1.  
DR PROSITE; PS50078; POLO\_BOX; 2.  
DR PROSITE; PSS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PSS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PSS00108; PROTEIN\_KINASE\_ST; 1.  
KW Transferrase, Serine/threonine-protein kinase; ATP-binding; Repeat;  
KW Phosphorylation; Alternative splicing.  
FT DOMAIN 63 315  
FT NP\_BIND 69 77  
FT BINDING 92 92  
FT ACT\_SITE 186 186  
FT DOMAIN 455 518  
FT DOMAIN 552 622  
FT VSAPLIC 373 373  
FT CONFLICT 386 386  
SQ SEQUENCE 631 AA; 70012 MW; 20857541870DBID2 CRC64;  
I -> V (IN REF. 2).  
/ftid=VSP\_004927.

Query Match	11.7%	Score 593;	DB 1;	Length 631;
Best Local Similarity	33.1%	Pred. No. 2,4e-24;		
Matches	122;	Conservative	64;	Mismatches 111; Indels 72; Gaps 4
QY	15	GNLLGKSPAGVYRAESIHITGLEVAIKMIDKKAMTKAGVORVONEVXIKHCOLKHPSTLE	74	
DB	66	GRLLGKGFARCYEATDRESGIAVAVKVIPOSRVAKPKQREKILNIEILHRDLOHRHIVR	125	
QY	75	LYNYPEDSNNYVYLVEWNGENBMRVYLKNRVKPESENBARPHMOIITGMVLYSHGILL	134	
DB	126	FSHFPEEDADNIYFIETELCSRKSLAHIMKAR-HTLEPEVRVYLRDITLGGKLYKHORGILL	184	
QY	135	RDLTLNMLLTRANNNIKIALDFGLATOLKMPHEKHVTLGTPNYISPEIATRSANGLSDV	194	
DB	185	RDLKLGNFPIIDNNELKVGDFGLARLEPPEQRKKTIGTPRYVAPEVYLRCGHGPEADV	244	
QY	195	WSLIGCMFPTLLIGRPPEPTDVYKNTLNKVYLVADYMPSEGLEADLIHQLRRHPARL	254	
DB	245	WSLCCVMATLLCGSPPEFTADLKETRYCRICKOVHYTLPASLSPAKQLALAIRASPRDRP	304	
QY	255	SLSSVLDPHFMNR-----	277	NSSTYSKDLG
DB	305	SIEQILIRHDFPTKGTVPRLPVSSCVTPVDPILPPNPARSLFAKYTKSLFGRKNNKXKHS	364	
QY	278	TVEESID-----GSHATISRAITASSSTISGSLFDRKRLLIQPLPNKMTVPFPKXSS	331	
DB	365	EDQDNVSCLAIPVSSQAPASLIETLAEDSSPRGTU-----	399	
QY	332	TDFSSSGDG 340		
DB	400	---ASSGDG 405		
RESULT 3				
PNK_RAT				

ID	CNR	RAT	STANDARD;	PRT;	615	AA.
AC	Q9R011;					
DT	16-OCT-2001	(Rel. 40, Created)				
DT	16-OCT-2001	(Rel. 40, Last sequence update)				
DT	16-OCT-2001	(Rel. 40, Last annotation update)				
DE	Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-inducible kinase) (Fragment).					
DE	CNK OR FAK.					
GN	Rattus norvegicus (Rat).					
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OC	NCBI_TaxID=10116;					
OK	[1]					
RN	SEQUENCE FROM N.A.					
RP	MEDLINE=99452760; PubMed=10523297;					
RX	Kanselmann G., Weller M., Mulf P., Jesberger S., Konietzko U., Scaffidi J., Staudl U., Bereliter-Hahn U., Strubhardt K., Kuhl D.;					
RA	"the polo-like protein kinases Fsk and Snk associate with a Ca(2+)-and integrin-binding protein and are regulated dynamically with synaptic plasticity.";					
RT	EMBO J. 18:5528-5539(1999).					
RL	-1 FUNCTION: Serine/threonine protein kinase involved in regulating M phase functions during the cell cycle. May also be part of the signaling network controlling cellular adhesion. In vitro, is able to phosphorylate CDC25C and casein (By similarity).					
CC	-1 CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.					
CC	-1 SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This interaction probably occurs via the POLO-box domain.					
CC	-1 SUBCELLULAR LOCATION: WHEN INDUCED, IT TRANSLOCATES INTO THE DENDRITES OF ACTIVATED NEURONS.					
CC	-1 TISSUE SPECIFICITY: Constitutively expressed in post-mitotic neurons.					
CC	-1 INDUCTION: By the intense actively associated with seizures.					
CC	-1 PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS CELLS EXIT MITOSIS (BY SIMILARITY).					
CC	-1 SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDC5/Polo subfamily.					
CC	-1 SIMILARITY: Contains 2 POLO box domains.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).					
CC	-----					
CC	EMBL; AF136584; AAF08367.1; -.					
DR	InterPro; IPR000959; POLO box.					
DR	InterPro; IPR000719; Prot_kinase.					
DR	InterPro; IPR008271; Ser_Thr_kin_AS.					
DR	InterPro; IPR002290; Ser_Thr_kinase.					
DR	Pfam; PF00069; Pkinase; 1.					
DR	Pfam; PF00659; POLO box; 2.					
DR	ProDom; PD000001; Prot_kinase; 1.					
DR	SMART; SM00220; S_TKc; 1.					
DR	PROSITE; PS50078; POLO_BOX_2.					
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP_1.					
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM_1.					
DR	PROSITE; PS00108; PROTEIN_KINASE_ST_1.					
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat; Phosphorylation.					
KW	NON_TER	1				
FT	DOMAIN	31	283	PROTEIN KINASE.		
FT	NP_BIND	37	45	ATP (BY SIMILARITY).		
FT	BINDING	60	60	ATP (BY SIMILARITY).		
FT	ACT_SITE	154	154	BY SIMILARITY.		
FT	DOMAIN	440	503	POLO BOX 1.		
FT	DOMAIN	537	607	POLO BOX 2.		
FT	NON_TER	615	615			
SQ	SEQUENCE	615 AA;	68800 MW;	06584C4229B5D71BD CRC64;		

CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.ebi.ac.uk/announcements">http://www.ebi.ac.uk/announcements</a>
CC	or send an email to <a href="mailto:license@ebi.ac.uk">license@ebi.ac.uk</a> ).
CC	-----
DR	EMBL; AF136584; AAP08367.1; -
DR	InterPro; IPR000959; Polo_box.
DR	InterPro; IPR000719; Prot_kinase.
DR	InterPro; IPR008271; Ser_thr_pkin_AS.
DR	InterPro; IPR002290; Ser_thr_pkinase.
DR	Pfam; PF00069; Kinase; 1.
DR	Pfam; PF00659; Polo_box; 2.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	SMART; SM00220; S_TKc_1.
DR	PROSITE; PS50078; POLO_BOX; 2.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KV	Phosphorylation.
FT	NON TER 1
FT	DOWAIN 31
FT	NP_BIND 37
FT	BINDING 60
FT	ACT_SITE 154
FT	DOMAIN 440
FT	FT DOMAIN 537
FT	NON TER 615
SQ	SEQUENCE 615 AA; 68800 MW; 06584C22B5D71BD CRC64;

Query Match      11.6%, Score 590.5; DB 1; Length 615;

Best Local Similarity 33.9%; Pred. No. 3.1e-24;  
Matches 126; Conservative 64; Mismatches 121; Indels 61; Gaps 7;

QY 15 GNLDKGFAGVYRAESHTGLVLAIKIKDKKAMKQVONVEYKIHCOLKHPSTLE 74  
DB 34 GRLLKGGFACCYETDTTETSIAYAVKVIPOSRAVKPHQREKIIINEIHLRDLQHRHIVR 93  
QY 75 LYNFEDSNVYVLYLEMCHGNEMRYLKNRYKPSSENAHMHQITTMGLYHSHQILH 134  
DB 94 FSHHEPDADNYITFELCSKSLAHMKAR-HTLEPEVRKYKQILSLGLKYLHQRGILH 152  
QY 135 RDLTSLNLLTRNNIKIADFGLATQLMPEKHYTLGTPNYISPEIATRSAGLESVD 194  
DB 153 RDLKGNFFITDNNMELKGFGLAARLEPQRKKTICGPNVAVPEVLARQHGPPADV 212  
QY 195 WSLGCMFETLLIGRPPTDPTVKNLKNVYADVEMSPSTIEAKDILHQLRRNPADRL 254  
DB 213 WSLGCMWTTLLGSPPEETADLKEITRCIKQVHTTLPASISLPARQLAAITLRASPPDRP 272  
QY 255 SLSSVLDHPMRSNSTKSKDGLTVEDSIDSGHATTISTATPASSST-----SIGG 304  
DB 273 SIEQILHDFEFTK-----GYTPRL-----PVSSCTTVEDLPPNPASLFAKVTK 318  
QY 305 SLPEPKR-----KLLIGQP-----LPNKMVFPKPKKST----- 332  
DB 319 SLFGRRKSKKKNHSEEDNVSLVSGIMRTSIGHPDVREPAPASALAPVLTETAEDS 378  
QY 333 -----DFSSSGDG 340  
DB 379 SPRGTIASSGDG 390

RESULT 4  
CNK\_HUMAN  
ID Q9H4B4; Q15767; STANDARD; PRT; 646 AA.  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FCF-inducible kinase) (Proliferation-related kinase).  
GN CNK OR FNK OR PRK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=20493044; PubMed=11039900;  
RA Kauselmann G., Rehl M., Andreesen R., Kautmann M., Kuhl D., Strubhardt K.;  
RT "Adhesion induced expression of the serine/threonine kinase Fnk in human macrophages.";  
RL Oncogene 19:4832-4839 (2000).  
RN [2]  
RP SEQUENCE OF 28-646 FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=96325053; PubMed=8702627;  
RA Li B., Ouyang B., Pan H., Reilemann P.T., Slamon D.J., Arceci R., Lu L., Dai W.;  
RT "Ptk, a cytokine-inducible human protein serine/threonine kinase whose expression appears to be down-regulated in lung carcinomas.";  
RN J. Biol. Chem. 271:19402-19408 (1996).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=96019242; PubMed=9353331;  
RA Ouyang B., Pan H., Lu L., Stambrook P., Li B., Dai W.;  
RT "Human Ptk is a conserved protein serine/threonine kinase involved in regulating M phase functions.";  
RN J. Biol. Chem. 272:28646-28651 (1997).  
RN [4]

RP CHARACTERIZATION.  
RX MEDLINE=20027391; PubMed=10557092;  
RA Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;  
RT "The physical association and phosphorylation of Cdc25C protein phosphatase by Ptk.";  
RL Oncogene 18:6029-6036 (1999).  
CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M phase functions during the cell cycle. May also be part of the signaling network controlling cellular adhesion. In vitro, is able to phosphorylate Cdc25C and casein.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This interaction probably occurs via the POLO-box domain.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- TISSUE SPECIFICITY: Transcripts are highly detected in placenta, lung, followed by skeletal muscle, heart, pancreas, ovaries and kidney and weakly detected in liver and brain. May have a short half-life. In cells of hematopoietic origin, strongly and exclusively detected in terminally differentiated macrophages. Transcript expression appears to be down-regulated in primary lung tumor.  
CC -1- INDUCTION: Cytokine and cellular adhesion trigger FNK induction.  
CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS CELLS EXIT MITOSIS (BY SIMILARITY).  
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDC5/Polo subfamily.  
CC -1- SIMILARITY: Contains 2 POLO box domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ293866; CAC10659.1; -;  
DR EMBL; U56998; AAC50637.1; ALT\_INT.  
DR Genbank; HGNC:2154; CNK.  
DR GK; Q9H4B4; -;  
DR MIM; 602913; -;  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.  
DR GO; GO:0000074; P:regulation of cell cycle; TAS.  
DR InterPro; IPR000959; POLO\_box.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00659; POLO\_box; 2.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PSS0078; POLO\_BOX; 2.  
DR PROSITE; PSS0107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; 1.  
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat; Phosphorylation.  
FT DOMAIN 62 314 PROTEIN KINASE.  
FT NP\_BIND 68 76 ATP (BY SIMILARITY).  
FT BINDING 91 91 ATP (BY SIMILARITY).  
FT ACT\_SITE 185 185 BY SIMILARITY.  
FT DOMAIN 470 537 POLO\_BOX 1.  
FT DOMAIN 567 637 POLO\_BOX 2.  
FT CONFLICT 99 99 V -> A (IN REF. 2).  
FT CONFLICT 353 353 H -> D (IN REF. 2).  
FT CONFLICT 419 419 H -> D (IN REF. 2).  
FT CONFLICT 464 470 FSEWGF -> VSKWVDY (IN REF. 2).  
FT CONFLICT 522 522 R -> P (IN REF. 2).  
SQ SEQUENCE 646 AA; 71789 MW; C20147CDDFC8A3B4 CRC64;

Query Match 11.6%; Score 589.5; DB 1; Length 646;  
Best Local Similarity 37.5%; Pred. No. 3.7e-24;

Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLGKGSFAGYVRAESHTGLEVAKIMDKAMRYAGVQVQVNEVYKHQCKHPSIE 74  
 DB 65 GRLGKGAFRCYEAADTETGSAVAKVI PQSRVVPKPHOREKLIETELHRLQHHIVR 124  
 QY 75 LNNYPEDSNVYVLYVEMCHNGEMNRYLKRVKVPFSSENEARHPHQIITGMVLYSHGILH 134  
 DB 125 FSHHEDADNITYIFELCSKRSIAHWKAR-HTLEPEVRYIRQLVSGIKYLHOGIILH 183  
 QY 135 RDLTSLNLLTTRNNMIKIDFGIATQKMPHEKHVTLCTPNYISPEIATRSAGLESVDV 194  
 DB 184 RDLKGNFPIENMELKVDGFLAARLEPDEQSKITICGTPNVVADEVLLRQHGGEADV 243  
 QY 195 NSLGCFFYTLIGRPPFDITVKATLANKVLYADYEMPSTLSIAKDLHQLRRNADRL 254  
 DB 244 NSLGCVMYTLIGSPFFETADIKETRCIKOVYITLPASISLPARQLLAIRASPRDRP 303  
 QY 255 SLSSVLDHPFMRNSTSKDGTVEDSIDSGHATISTAVTSSST-----SISG 304  
 DB 304 SLDQILRHDPFTR-----GYTPDRL-----PISCVTVPDLTTPNPASLPAKATV 349  
 QY 305 SLFDRKR 311  
 DB 350 SLFVRKK 356

RESULT 5  
 SNK\_MOUSE STANDARD; PRT; 682 AA.

ID SNK\_MOUSE  
 AC P5351;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/chreonine-protein kinase SNK (EC 2.7.1.-) (Serum inducible kinase).  
 GN SNK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92375085; PubMed=1508211;  
 RA Simmons D.L., Neel B.G., Stevens R., Ewelt G., Ertkeon R.L.;  
 RT "Identification of an early-growth-response gene encoding a novel putative protein kinase."  
 RT Mol. Cell. Biol. 12:4164-4169(1992).  
 RL -1- FUNCTION: May play a role in the division of at least some cell types, such as fibroblasts, and could function in embryogenesis, wound healing or neoplasia.  
 CC -1- TISSUE SPECIFICITY: Brain, lung and heart.  
 CC -1- INDUCTION: By serum and phorbol ester.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDCs/Polo subfamily.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL; M96163; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A4493; A4493.  
 DR MGD; MGI:1099780; Snk.  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; I.

DR Pfam; PF00659; POLO\_box; 2.  
 DR Prodom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS50078; POLO\_BOX; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferase; Serine/chreonine-protein kinase; ATP-binding; Repeat.  
 FT DOMAIN 54 59 POLY-HIS.  
 FT BINDING 79 331 PROTEIN\_KINASE.  
 FT NP\_BIND 85 93 ATP (BY SIMILARITY).  
 FT ACT\_SITE 108 108 ATP (BY SIMILARITY).  
 FT BINDING 202 202 BY SIMILARITY.  
 FT DOMAIN 507 570 POLO\_BOX\_1.  
 FT DOMAIN 603 674 POLO\_BOX\_2.  
 SQ SEQUENCE 682 AA; 77811 MW; 586DEADF208A9D CRC64;

Query Match 11.1%; Score 561.5; DB 1; Length 682;  
 Best local similarity 41.6%; Pred. No. 1.2e-22;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLGKGSFAGYVRAESHTGLEVAKIMDKAMRYAGVQVQVNEVYKHQCKHPSIE 74  
 DB 82 GYVLGKGFACVCEMTDLTNNKRYVAKIIPHSRVKPHOREKIDKEIELHRLHKGAVQ 141  
 QY 75 LNNYPEDSNVYVLYVEMCHNGEMNRYLKRVKVPFSSENEARHPHQIITGMVLYSHGILH 134  
 DB 142 FYHVEDKKNITYILCYCSRRSMHILKAR-KVLEPEVRYIRQLVSGIKYLHOGIILH 200  
 QY 135 RDLTSLNLLTTRNNMIKIDFGIATQKMPHEKHVTLCTPNYISPEIATRSAGLESVDV 194  
 DB 201 RDLKGNFPIENMELKVDGFLAARLEPDEHRRRTICGTPNVVADEVLLRQHGGEADI 260  
 QY 195 NSLGCFFYTLIGRPPFDITVKATLANKVLYADYEMPSTLSIAKDLHQLRRNADRL 254  
 DB 261 NALGCVMYTMLGRPPFTTNKETYRCIRARYTPSSILAPAKLILASMLSKNEDRP 320

QY 255 SLSSVLDHPF 264  
 DB 321 SLDDIIRHDF 330

RESULT 6  
 SNK\_RAT STANDARD; PRT; 682 AA.

ID SNK\_RAT  
 AC Q9R012;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/chreonine-protein kinase SNK (EC 2.7.1.-) (Serum inducible kinase).  
 GN SNK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9452760; PubMed=10523297;  
 RA Kausselmann G., Weller M., Wulff P., Jessberger S., Konietzko U., Scafield J., Straubli U., Berleiter-Hahn J., Streibhardt K., Kuhl D.;  
 RT "The polo-like protein kinases Pnk and Snk associate with a Ca(2+)-and integrin-binding protein and are regulated dynamically with synaptic plasticity."  
 RT EMBO J. 18:5528-5539(1999).  
 RL -1- FUNCTION: May play a role in the division of at least some cell types, such as fibroblasts, and could function in embryogenesis, wound healing or neoplasia. (By similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF136583; AAF08366.1; -  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00659; POLO\_box; 2.  
 DR Prodom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS50078; POLO\_BOX; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat.  
 KW DOMAIN 54 59 POLY-HIS.  
 FT DOMAIN 79 331 PROTEIN\_KINASE.  
 FT NP\_BIND 85 93 ATP (BY SIMILARITY).  
 FT BINDING 108 108 ATP (BY SIMILARITY).  
 FT ACT\_SITE 202 202 BY SIMILARITY.  
 FT DOMAIN 570 570 POLO\_BOX 1.  
 FT DOMAIN 603 674 POLO\_BOX 2.  
 SQ SEQUENCE 682 AA; 77919 MW; 58C50DBDE83D5F3 CRC64;

Query Match 11.0%; Score 560.5; DB 1; Length 682;  
 Best Local Similarity 41.6%; Pred. No. 1.4e-22;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLGKGSFAGYVRAHSHTGLEVATKMDKKAMYGAVQVQVNVKTHCQKHSILE 74  
 Db 82 GVLGSGGFGKCEMTDLNNKYVAKTTPSRVAPRQREKIDKEIRLHMHVVO 141  
 QY 75 LYNVEDSNVYVLVLEMCHNGENNRVYKRVKPFSENEARHFMHQITLMLYLSHGILH 134  
 Db 142 FHYHEDDKENITYLEYGRSMAMHLKAR-KVLEPEVRYVLRQVSGLKXILHBEELH 200  
 QY 135 RLUTLSNLLTTRWNTKIKADFGLATOLKMPHEKHYTLCTPNYISPELATRSAGLESVP 194  
 Db 201 RLKIKGNPFINAMEIKVDFGLAARLEBIEHRRRTICGTPYVLSPEVANKGHCESPI 260  
 QY 195 WSIGMCFYLLIGRPFDYDVTYKNTLKVLYADYEPSFLSTLEADLIHQILRRPADRL 254  
 Db 261 WALGCVMTLWLGRRPFETTLNKEVTRCIRREARYTPSSLAPAKHLIASMLSKPEDRP 320  
 QY 255 SLSSVLDHPF 264  
 Db 321 SLDDIIRHDF 330

RESULT 7  
 ID SNK\_HUMAN STANDARD; PRT; 685 AA.  
 AC Q9NXY3; O60679; O96CV7; O9UE61;  
 DT 16-OCT-2001 (Rel. 40. Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase SNK (BC 2.7.1.-) (Serum inducible  
 DE kinase).  
 GN SNK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN MAMMALIA  
 RP SEQUENCE FROM N.A.  
 RA Ouyang B., Dai W.,  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=colon.  
 RA Anderson K.M., Nerurkar S.S., Hansbury M.J., Formwald J., Scott G.,  
 RA Bouvry K., Mui P., Imbruglia C.S., Carlson K., Marshall L.A.,  
 RA Roshak A.K.;  
 RT "Identification and characterization of human serum-inducible Kinase  
 RT (SNK), a novel member of the polo-kinase family of cell cycle  
 RT regulators: potential implication for regulation of vascular smooth  
 RT muscle proliferation.";  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=skin;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Tshibuyi S., Cantucci P., Prange C.,  
 RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 110-408 FROM N.A.  
 RA Fidler C., Boulwood J., Wang Jabs E., Waincoat J.S.;  
 RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: May play a role in the division of at least some cell  
 CC types, such as fibroblasts, and could function in embryogenesis,  
 CC wound healing or neoplasia (By similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC FT DOMAIN 82 334 PROTEIN KINASE.
CC FT NP_BIND 88 96 ATP (BY SIMILARITY).
CC FT BINDING 111 111 ATP (BY SIMILARITY).
CC FT ACT_SITE 205 205 BY SIMILARITY.
CC FT DOMAIN 510 573 POLO BOX 1.
CC FT DOMAIN 606 677 POLO BOX 2.
CC FT CONFLICT 28 28 A -> G (IN REF. 1).
CC SEQ SEQUENCE 685 AA; 78236 MW; 6429P6F8D30B33 CRC64;

Query Match 11.0%; Score 560.5; DB 1; Length 685;
Best Local Similarity 41.6%; Pred. No. 1,4e-22;
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLGKGFAGYVRAESIHTEGLVALIKMDKKAMKAGVQRQNVKIKHCOLKHSITL 74
DB 85 GVLGKGFAGYVRAESIHTEGLVALIKMDKKAMKAGVQRQNVKIKHCOLKHSITL 144
QY 75 LYNVEDSNVYVLEEMCHNGEMNRYLKRVKPFSENEARHFMHQITITGLYLHSHGILH 134
DB 145 FHYVEDKENIYILLEYCSRSMHILKAR-KVLTEPRVYVRLQIVSGIKYHGEILH 203
QY 135 RDLVLSNLLITNNMIKIDFGLATQKMPHEKHYTLGTPNYISPEIATRSAGLESVD 194
DB 204 RDLKGNFPIINAMELKVDGFLARLEPLEHRRRTICGTPNYLSPEVINKQGHGCSDI 263
QY 195 NSLGMFVYLLIGRPFEDTDVTKNTLVADYEMPSFLISAKDLIHQLLRNADRL 254
DB 264 WMLGCMYVTLGRPFETTNLKEVYRCIRKARYMPSLLAPAKHLIASMSKNEPDR 323
QY 255 SLSSVLDHPF 264
DB 324 SLDDIRHDF 333

RESULT 8
STR6_XENLA STANDARD; PRT; 407 AA.
ID STR6_XENLA
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase Eg2 (BC 2.7.1.37) (p46Eg265).
GN Eg2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxId=8355;
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RC MEDLINE=98122855; PubMed=9454730;
RA Roghi C., Giet R., Uzbekov R., Morin N., Chartreain I., Le Guellec R.,
RA Couturier A., Doree M., Philippe M., Prigent C.;
RT "The Xenopus protein kinase Eg2 associates with the centrosome in a
RT cell cycle-dependent manner, binds to the spindle microtubules and is
RT involved in bipolar mitotic spindle assembly."
RT J. Cell Sci. 111:557-572(1998).
CC -1- FUNCTION: Associates with the centrosome in a cell-cycle dependent
CC manner and invades the microtubules at the poles of the spindle
CC during mitosis suggesting that it may be involved in the correct
CC formation of bipolar mitotic spindles.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC especially from prophase through anaphase. Partially colocalised
CC with gamma tubulin in the centrosome, from S to M phase.
CC -1- TISSUE SPECIFICITY: Highly expressed in ovary and testis.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-eb.ch/announce/
CC or send an email to license@isb-eb.ch).
CC -----
CC EMBL; 217207; CAA78915.1; -.
CC PIR; S52243; S52243.
CC DR HSSP; P24941; 1A01.
CC DR InterPro; IPR000719; Prot_kinase.
CC DR InterPro; IPR008271; Ser_thr_kin_AS.
CC DR InterPro; IPR002290; Ser_thr_kinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR Cell cycle; Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation.
CC FT DOMAIN 140 390 PROTEIN KINASE.
CC FT NP_BIND 146 154 ATP (BY SIMILARITY).
CC FT BINDING 169 169 ATP (BY SIMILARITY).
CC FT ACT_SITE 263 263 BY SIMILARITY.
CC SEQ SEQUENCE 407 AA; 46372 MW; DE1628A2C6D11277 CRC64;

Query Match 10.9%; Score 554.5; DB 1; Length 407;
Best Local Similarity 39.7%; Pred. No. 1,4e-22;
Matches 106; Conservative 60; Mismatches 98; Indels 3; Gaps 2;

QY 9 IEDFKVGNLKGSGFAGYVRAESIHTEGLVALIKMDKKAMKAGVQRQNVKIKHCOLK 68
DB 137 LEDFEIGRLGKGFAGYVRAESIHTEGLVALIKMDKKAMKAGVQRQNVKIKHCOLK 196
QY 69 HPSILELYNVEDSNVYVLEEMCHNGEMNRYLKRVKPFSENEARHFMHQITITGLYLH 128
DB 197 HPVILRLYGVFHDASHRVYLLDYAPGSELFRLEQKCTR-FDDQRSAMYIKQLAEALLYCH 255
QY 129 SHGIIHRDLTSLNLLITNNMIKIDFGLATQKMPHEKHYTLGTPNYISPEIATRSAGH 188
DB 256 SKKVHNRDIPKPELTLGSGNGELKIDFGWS--VHABSSRTTLGTLDPFPMIGRW 313
QY 189 GLESDVNSLGMFVYLLIGRPFEDTDVTKNTLVADYEMPSFLISAKDLIHQLLRNADRL 248
DB 314 DEIVDMSLGVLCYELVQKPFETDTHTQETRRISKVFGQRPVYSEARLDVSLK 373
QY 249 NPADRLSLSSVLDHPFMSNNSSTKSD 275
DB 374 NPVHRLPLKGVLEHPWIIKNSQLKKD 400

RESULT 9
STR6_XENLA STANDARD; PRT; 408 AA.
ID STR6_XENLA
AC 091819;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase Eg2-like (BC 2.7.1.37) (p46XLEg22).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxId=8355;
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RC MEDLINE=98122855; PubMed=9454730;
RA Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;
RT "Eg2, selected by differential screening encodes a new Xenopus protein
RT kinase family."
RT Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis

```

CC	especially from prophase through anaphase. Partially colocalised with gamma tubulin in the centrosome, from S to M phase (By similarity).
CC	-1- PTM: Phosphorylated (By similarity).
CC	-1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. Aurora subfamily.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; Z17206; CAA18914.1; ALT_INTT.
DR	HSSP; P24941; LAQ1.
DR	InterPro; IPR000719; Prot Kinase.
DR	InterPro; IPR008271; Ser Thr pkin AS.
DR	InterPro; IPR002290; Ser thr kinase.
DR	InterPro; IPR001245; Tyr pkinase.
DR	Pfam; PF00069; pkinase; 1.
DR	ProDom; PD000001; Prot_Kinase; 1.
DR	SMART; SM00220; S_TKc; 1.
DR	SMART; SM00219; TyrcK; 1.
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR	PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW	Cell cycle; Transferase; Serine/threonine-protein kinase; ATP-binding; Phosphorylation.
FT	DOMAIN 140 390 PROTEIN KINASE.
FT	NP_BIND 146 154 ATP (BY SIMILARITY).
FT	BINDING 169 169 ATP (BY SIMILARITY).
FT	ACT_SITE 263 BY SIMILARITY.
SQ	SEQUENCE 408 AA; 46461 MW; 976A69C7357AE8 CnC64;
Query Match:	10.8%; Score 548; DB 1; Length 408;
Best Local Similarity	39.9%; Pred. No. 3.2e-22;
Matches 107;	Conservative 60; Mismatches 97; Indels 4; Gaps 3;
QY	9 IEDPKVNLGKSGPAGYRAESIHTGQLEVAIKMIDKKAMYKAGVORVQNEVKTHOLK 68
DB	137 LEDFPIGRPLGKRGKGVYLLARBRSKRTIALNLKYLFSQLEKAGVHQREVEIQSHR 196
QY	69 HPSIIELNYPEDSNVYVYLVEIMCNGEMNRYLKNRVKPFSENEARRHFMQIITGMYLH 128
DB	197 HPNIIRLKGYPHDASRYVLLIDYAPGGLFRELQKCTR-FDDORSALYIKQLAELLALYCH 255
QY	129 SHGILHRDLTSLNLLTNNNNIKIADFLATQLMPEHKHYTLGTPNYISPELATSAH 188
DB	256 SKVYIHRDKRENLLSGNGELKIDRWMS--VIAFSSRRITTLGGLDYLRPEMTGRMH 313
QY	189 GLESDVSLGCMFTYLLIGRPPEPTDVTYKNTLNKRVLADYEMPSFLSIEAKDLIHOILRR 248
DB	314 DEKVDLWSLGLCYEFLVIGKPEPTDTHQETRYRISKVEFGYPYVSEAKDLVSKLLKH 373
QY	249 NPADRLSLSYLDHPMSRNS-STKSD 275
DB	374 NPNHRLPLKGLVLEHFMVYKNSQPEKKD 401
RESULT 10	
PLK1 MOUSE	
ID	PLK1 MOUSE STANDARD; PRT; 603 AA.
AC	007832;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Serine/threonine-protein kinase PLK (EC 2.7.1.1-) (PLK-1) (Serine-threonine protein kinase 13) (STPK13).
GN	Mus musculus (Mouse).
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Bone marrow;  
 RA MEDLINE=9328160; PubMed=8099445;  
 RT Clay F.J., McEwen S.J., Bertoneello I., Wilks A.F., Dunn A.R.;  
 RT "Identification and cloning of a protein kinase-encoding mouse gene,  
 RT Plk, related to the polo gene of Drosophila.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:4882-4886(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 x CBA; TISSUE=Thymus;  
 RA MEDLINE=9428923; PubMed=8018577;  
 RT Hamanaka R., Malold S., Smith M.R., O'Connell C.D., Longo D.L.,  
 RA Ferris D.K.;  
 RT "Cloning and characterization of human and murine homologues of the  
 RT Drosophila polo serine-threonine kinase.";  
 RL Cell Growth Differ. 5:249-257(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=94067140; PubMed=7902533;  
 RA Lake R.J., Jelinek W.R.;  
 RT "Cell cycle- and terminal differentiation-associated regulation of  
 RT the mouse mRNA encoding a conserved mitotic protein kinase";  
 RL Mol. Cell. Biol. 13:7793-7801(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marnettin K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rata S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hilyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Roderiguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RC [1]- FUNCTION: May be required for cell division and may have a role  
 during G1 or S phase.  
 CC [2]- SUBCELLULAR LOCATION: Nuclear.  
 CC [3]- TISSUE SPECIFICITY: Newborn and adult spleen, fetal and newborn  
 kidney, liver, brain, thymus and adult bone marrow, thymus,  
 CC ovary and testes.  
 CC [4]- DEVELOPMENTAL STAGE: In the thymus, levels increased during fetal  
 development, were highest in newborn animals and decreased in the  
 CC adult. In the testes, the PLK levels were higher in the adult than  
 CC in prepubescent mice while in the ovary, the levels were higher in  
 CC the prepubescent mice. Accumulates to a maximum during the G2 and  
 CC M phases, declines to a nearly undetectable level following  
 CC mitosis and throughout G1 phase, and then begins to accumulate  
 CC again during S phase.  
 CC [5]- SIMILARITY: Belongs to the ser/thr family of protein kinases.  
 CC [6]- SIMILARITY: Contains 2 POLO box domains.  
 CC [7]- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L06144; AAA59948.1; -  
 DR EMBL; U01063; AAA56635.1; -  
 DR EMBL; L19558; AAA16071.1; -  
 DR EMBL; BC006880; AAA06880.1; -  
 DR PIR; A47545; A47545.  
 DR PIR; A45961; A4596.  
 DR MGD; MGI:97621; PLK.  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00659; PKinase; 1.  
 DR Pfam; PF00659; POLO\_box; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KW Nuclear protein.  
 FT DOMAIN 53 305 PROTEIN\_KINASE.  
 FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
 FT BINDING 82 82 ATP (BY SIMILARITY).  
 FT ACT\_SITE 176 176 BY SIMILARITY.  
 FT DOMAIN 417 480 POLO\_BOX 1.  
 FT DOMAIN 515 584 POLO\_BOX 2.  
 FT CONFLICT 4 4 A -> V (IN REF. 1).  
 FT CONFLICT 15 15 A -> T (IN REF. 1).  
 FT CONFLICT 23 23 P -> L (IN REF. 1).  
 FT CONFLICT 27 27 P -> A (IN REF. 1).  
 FT CONFLICT 29 29 G -> S (IN REF. 1).  
 FT CONFLICT 41 41 P -> L (IN REF. 1).  
 FT CONFLICT 54 54 V -> I (IN REF. 1).  
 FT CONFLICT 495 495 A -> R (IN REF. 1).  
 FT SEQUENCE 603 AA; 68300 MW; 1B980646366EFA10 CRC64;

Query Match 10.7%; Score 545; DB 1; Length 603;  
 Best Local Similarity 27.2%; Pred. No. 7, 8e-22;  
 Matches 152; Conservative 95; Mismatches 190; Indels 122; Gaps 20;

QY 15 GNLLGKGSFAGVYRAESIHTGLEVALIKMIDKKMYRAGVQVONEVKIHCQKPSILE 74  
 DB 56 GRFLGKGSFAGVYRAESIHTGLEVALIKMIDKKMYRAGVQVONEVKIHCQKPSILE 115  
 QY 75 LKMYFEDSNVYIVLMECHNGENKRYLKRVPFSENEARHPHQIITGMLYHSHGILH 134  
 DB 116 FHDFFDSDFVFAVLELCRRRSLLBLHKR-KALTEPEAAVYRQIVLGCQYHRRQVH 174  
 QY 135 RDLTSLNLLTRNMKIKADFGLATQKMPHEKHYLTGPNYISPEIATRSNAGLESIV 194  
 DB 175 RDLKLGNTLNLBLEVKIGDFGATATVVEGEKKTLCGPNYIAEVLVSKSHSEVDV 234  
 QY 195 WSLGCFYTLTIGRPFDTDVTNKLNVLADYEMPSFLISBAKOLIHOLLRRNPADRL 254  
 DB 235 WSGICIMYTLVQKPPFETSCLEKRYIKRNEYSPKHNIPVAASLIQMLQTDPTAP 294  
 QY 255 SLSSVLDHFMGNSSTKSKDCTVEDSIDSGHATIS---TATASSSTISGSLPD--K 309  
 DB 295 TLHLLNDEFT-----SGYIPARLPITCLTIPRPSIASSIDPSS 336  
 QY 310 RLL-----IGPLPNKMTVPFKNKSSTDFSSSGDG-----NSFYTQMGQETSQ-SGR 357  
 DB 337 RKLKLVKLVNGVNEPLDR-----PREKEPVRRTNNAIECHLSDDLQQLTVNASKPSER 392  
 QY 358 GRVYQDAEERPH-----SRVLRAYSSDRSGTS-----NSQSAKTYTMEER 398  
 DB 393 GLVROEADDPACIPFVWSKWD---YSDKYGLGYQLCDNSVGVLFNDSTRILTY---- 445

QY 399 CHSAEMLSVSKSGGGENEERYSPDNNAN-----INPFKEKTS---SSGSPERPD 448  
 DB 446 -NDGDSLOYIERDG-----TESVLTWSHNSLMKKTLLNRYRWSEHLKAGANITPR 500  
 QY 449 NNQALSNHLCPGKTPPPDPPTPQETTVQOMF-----GMLQIN----- 486  
 DB 501 EGDLELA-----RLPY-----LRTWFRTSAILHLNSGVQINFPDHTKIL 543  
 QY 487 AHRKTEYDSISPNDFQ 505  
 DB 544 CPLMAAVTY--INERKDFQ 560

RESULT 11  
 PLK1\_RAT  
 ID PLK1\_RAT STANDARD; PRT; 603 AA.  
 AC 062673;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase PLK (EC 2.7.1.-) (PLK-1).  
 GN PLK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreeas;  
 RA Amstrup J., Hansen J.A., Hxirils Nielsen J.;  
 Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: May be required for cell division and may have a role  
 CC during G1 or S phase (By similarity).  
 CC - SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC - SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDC5/Polo subfamily.  
 CC - SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U0188; AAA1885.1; -  
 DR InterPro; IPR000959; POLO\_Box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00659; PKinase; 1.  
 DR Pfam; PF00659; POLO\_Box; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KW Nuclear protein.  
 FT DOMAIN 53 305 PROTEIN\_KINASE.  
 FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
 FT BINDING 82 82 ATP (BY SIMILARITY).  
 FT ACT\_SITE 176 176 BY SIMILARITY.  
 FT DOMAIN 417 480 POLO\_BOX 1.  
 FT DOMAIN 515 584 POLO\_BOX 2.  
 FT SEQUENCE 603 AA; 1074FEB3B7EDC002 CRC64;

Query Match 10.6%; Score 537; DB 1; Length 603;  
 Best Local Similarity 27.0%; Pred. No. 2, 1e-21;  
 Matches 151; Conservative 95; Mismatches 191; Indels 122; Gaps 20;

QY 15 GNLLGKSGFAGVYRAESHTTGLEVALIKMIDKKAMKAGVORVQNEVKIKCOLKHPSTLE 74  
 DB 56 GAFPLCKGGKACFELEISDPTKEVPFGKIVPKSLKPKHOKKMSMETSITHSLHONVVG 115  
 QY 75 LYNVEDSNVYVLLVEMCHNGEMNRYLKNRVKPFSENEARHPHQIITGMILYHSHGILH 134  
 DB 116 FHGFEDSDDFVAVVLELCRRSRSLLEIHKRR-KALTEPARAYYLQIVLGCQYLHNVYH 174  
 QY 135 RDLTSLNLLTRNNMIKTLADFGALATOLKMPHEKHVTLCTGPNYISPELATRSANGLESVD 194  
 DB 175 RDLKLGKGNFLNEDLEVKIGDFGLATKAYBGERKKTLCGTNYTALPEVLSKKGHSFEVDV 234  
 QY 195 WSLGCMFTLLIGRPFDPTVTKTLANKVVLADYEMSFSLSEAKDLIHOILRRNPADRL 254  
 DB 235 WSLGIMVTLVVGKRPFEFSCLEKETYLRKKNEXSIPGHNPVNASILQCKLQDPAPAR 294  
 QY 255 SLSSVLDHPFMSRNSSTKSNOLGVEDSIDSGHATIS---TATYASSSTISGSLFD-K 309  
 DB 295 TIHHLNDEFPT---SGYIPARLPITCLTIPRPSIAPSSIDPSN 336  
 QY 310 RRLI-----IGPLPNKMTVPFNKKSSTDFFSSGSD-----NSFYQMGNOETSN-SGR 357  
 DB 337 RKPFLTVANKGVENLDPDR---PREKEPVRETNMAIECHSDLIQOLTSVNAKPSER 392  
 QY 358 GRVYODAEERPH-----SRYLRAVSSDRSGTS-----NSQSQAKTYTMR 398  
 DB 393 GLVROEAEADPACIPFVWSKWD---YSDKYGLGYOLCDNSVGLFNDSTRILY----- 445  
 QY 399 CHSNAEMLSVKRSGCGENEREYPTDNAN-----INPEFEXTSS---SGSFERPD 448  
 DB 446 -NNGDSLOYIEROG---TESYLTVSHRPSLAKKTLTLANFVNWSEHLKAGANITPR 500  
 QY 449 NNQALSNHLCPGKTPFPFADPTQETVQWF-----GNLQIN----- 486  
 DB 501 EGDELA-----RLPY-----LMTWERTSAGIILHLNSGTQVQINFPQDHTKLIR 543  
 QY 487 AHUKRTTEYDISPNRDFQ 505  
 DB 544 GPLMAAVTY--INERKDFR 560

RESULT 12  
 PLK1 HUMAN  
 ID PLK1 HUMAN STANDARD; PRT; 603 AA.  
 AC P53350/  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serine/threonine-protein kinase PLK (EC 2.7.1.1-) (PLK-1) (Serine-  
 threonine protein kinase 13) (STPK13).  
 GN PLK OR PLK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94289293; PubMed=8018557;  
 RA Hamanaka R., Maloid S., Smith M.R., O'Connell C.D., Longo D.L.,  
 RA Ferris D.K.;  
 RT "Cloning and characterization of human and murine homologues of the  
 RL Drosophila polo serine-threonine kinase.";  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94067140; PubMed=7902533;  
 RA Lake R.J., Jelinek W.R.;  
 RT "Cell cycle- and terminal differentiation-associated regulation of  
 RL the mouse mRNA encoding a conserved mitotic protein kinase.";  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RT Mol. Cell. Biol. 13:7793-7801(1993).

RX MEDLINE=95051109; PubMed=7962193;  
 RA Golubev R.M., Schultz S.J., Bartek J., Ziemlecki A., Ried T.,  
 RA Nigg E.A.;  
 RT "Cell cycle analysis and chromosomal localization of human Plk1, a  
 RT putative homologue of the mitotic kinases Drosophila polo and  
 RT Saccharomyces cerevisiae Cdc5.";  
 RL J. Cell Sci. 107:1509-1517(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=94173904; PubMed=8127874;  
 RA Holtrich U., Wolf G., Brauning A., Karn T., Boehme B.,  
 RA Ruebman-Waigmann H., Streibhardt K.;  
 RT "Induction and down-regulation of PLK, a human serine/threonine  
 RT kinase expressed in proliferating cells and tumors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1736-1740(1994).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon, and Lung;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien P.,  
 RA Diachenko L., Marnett K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP PHOSPHORYLATION OF SER-335.  
 RX MEDLINE=22329537; PubMed=12442251;  
 RA Wind M., Kelm O., Nigg E.A., Lehmann W.D.;  
 RT "Identification of phosphorylation sites in the polo-like kinases  
 RT Plk1 and Plk1 by a novel strategy based on element and electrospray  
 RT high resolution mass spectrometry.";  
 RL Proteomics 2:1516-1523(2002).  
 CC -1- FUNCTION: May be required for cell division and may have a role  
 CC during G1 or S phase.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: Placenta and colon.  
 CC -1- DEVELOPMENTAL STAGE: Accumulates to a maximum during the G2 and M  
 CC phases, declines to a nearly undetectable level following mitosis  
 CC and throughout G1 phase, and then begins to accumulate again  
 CC during S phase.  
 CC -1- INDUCTION: By growth-stimulating agents.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDC5/Polo subfamily.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 DR EMBL; U01038; AAA56634.1; -;  
 DR EMBL; I19559; AAA36659.1; -;  
 DR EMBL; X73458; CAA51837.1; -;  
 DR EMBL; X75932; CAA53536.1; -;.

```

DR EMBL; BC002369; AAH02369.1; -.
DR EMBL; BC003002; AAH03002.1; -.
DR EMBL; BC014846; AAH14846.1; -.
DR PIR; S34130; S34130.
DR Genew; HGNC:9077; PLK.
DR GK; P53350; -.
DR MIM; 602098; -.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TMS.
DR GO; GO:0008283; P:cell proliferation; TMS.
DR GO; GO:0007067; P:mitosis; TMS.
DR GO; GO:0007074; P:regulation of cell cycle; TMS.
DR InterPro; IPR000959; POLO box.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00659; POLO box; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00078; POLO BOX; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW Nuclear protein; Phosphorylation.
FT DOMAIN 53 305
FT NP_BIND 59 67
FT BINDING 82 82
FT ACT_SITE 176 176
FT DOMAIN 417 480
FT DOMAIN 515 584
FT MOD_RES 335 335
FT CONFLICT 2 2
FT CONFLICT 11 11
FT CONFLICT 58 58
FT CONFLICT 60 60
FT CONFLICT 73 73
FT CONFLICT 141 141
FT CONFLICT 227 227
FT CONFLICT 301 301
FT CONFLICT 495 495
FT CONFLICT 501 501
SQ SEQUENCE 603 AA; 68254 MW; 178C2F13C10B8206 CRC64;

Query Match 10.6%; Score 536; DB 1; Length 603;
Best Local Similarity 32.2%; Pred. No. 2.3e-21;
Matches 118; Conservative 73; Mismatches 144; Indels 32; Gaps 6;

15 GNULGSGFAGVYRAHSIHGLEVAIKMIDKKMYKAGWQVRQNEVKIKHCOLKHPSTLE 74
56 GRIPLGSGFAGVYRAHSIHGLEVAIKMIDKKMYKAGWQVRQNEVKIKHCOLKHPSTLE 115
75 LNYVEDSNVYVLEMECHNGENRMYLKNRVKPFSENEARHFMHQITITGMLYLSHGILH 134
116 HGFPEPDNPFVFLVELCRRSRSLBLHKKR-KALTEPARVYVRLQVLGGQYVHRRVYH 174
135 RLDTLSNLLTRNNMTKIADFGIATQKPHKEKHYLLCGTPNTISPEIATRSAGHSVDY 194
175 RDLCKGNLFLNEDLEVKIGDFGLATKVEYGEKSKTKLCTGPNYIADEVLSKKGHSFEVY 234
195 WSIGCMFYLLTIGRPPEDTDVTNKLKYLADYEMPSFLSIKXOLHQLLRNPADRL 254
235 WSIGCMFYLLTIGRPPEDTDVTNKLKYLADYEMPSFLSIKXOLHQLLRNPADRL 294
255 SLSSVLDHFMNSSTKSKDCTVEDSIDSGHATIS--TATTAASSTISGSLPD-- 308
295 TINELENDPEPT-----SGVTPARLPITGLTIPRPSIAPSSLDPSN 336
309 -KRRLILGQPLPKMTVPFPAKSTSTPSSSGD-----NSFYQKNGQESN-SGGGRVY 361
337 RKPLTLYLNGLEBNLPERPREKEEPVVRTEGVVDCHLSQMLQOLHVSNAKPSERGLVR 396
362 QDAERP 368

```

```

DB 397 QDAERP 403

RESULT 13
ID KKK1 YEAST STANDARD; PRT; 1518 AA.
AC P34214;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (BC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288C;
RX MEDLINE=94078677; PubMed=8256524;
RA Paller C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RA Bolotin-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MBI gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT kinases."
RL Yeast 9:1149-1155(1993).
CC -1- Similarity: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
CC EMBL; X71133; CAA50456.1; -.
CC EMBL; Z28101; CAA81941.1; -.
CC PIR; S37928; S37928.
CC HSSP; 063450; 1A06.
CC GerMOnline; 139857; -.
DR SCD; S0001584; HSL1.
DR GO; GO:0005935; C:bud neck; IDA.
DR GO; GO:0005940; C:septin ring; IDA.
DR GO; GO:0004672; P:protein kinase activity; IDA.
DR GO; GO:000086; P:G2/M transition of mitotic cell cycle; IGI.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:000074; P:regulation of cell cycle; IMP.
DR GO; GO:000135; P:septin checkpoint; IGI.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 81 369
FT NP_BIND 87 95
FT BINDING 110 110
FT ACT_SITE 239 239
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 10.4%; Score 526.5; DB 1; Length 1518;
Best Local Similarity 22.3%; Pred. No. 2.5e-20;

```

Matches 253; Conservative 182; Mismatches 429; Indels 271; Gaps 42;

QY 7 EKLEPKVGNLKGKSPAGVYAAEIIHGLEVAIMDKA-----MYKGM 53  
 Db 76 DIVGVWKLKTKTGKSGCRRLAKMETGOLAATIVPKKAFVHCNNNGTVPNSYSSSM 135  
 QY 54 VOR-----VONEVKIHCQKPSIILELYNEPDSNYVLYLEM 91  
 Db 136 VTSNVSPSIASREHNSHQTNPGVIGIREIYIMGLISTNNMALFEVWENSELYLVLEY 195  
 QY 92 CHNGEMNNYLNKRVKVPSENEARHMHQIITGMILYHSHGILHDLTLNILL-TRNNMI 150  
 Db 196 VDGELFYLVSFKG-LPEREAHYFKQIVGVSVCHSFNCHDLPKRENLILKKORRI 254  
 QY 151 KIADFGATOLKMPHEKHYTLCGTPNYSPPIAT-PSAHGLESVWISGCMFTYLLGRP 209  
 Db 255 KIADFGMAA-LELPFKLLKTKSGSPHVASPEIIVWGRPHGSPSVWSCGIVLEFLLTGH 313  
 QY 210 PDDDTVNTLVNKLAVLDYEMPSFLISBANDLHOLLRRPADRLSLSVLDHPF----- 264  
 Db 314 PFNDNTIKQLLKVQSGKYQMPNSLSEARDLSIKLIVIPKRIITQOILKHPILKYD 373  
 QY 265 -----MSRSTKSKD-----LGTYESIDSGHA--TISTATASST--SIS 303  
 Db 374 DLPVNVKVLKRRKONMARGKNSDLHANNVSPITVHSGEIDSLRSLQILMHGVS 433  
 QY 304 GSLFDKRLILGOLPNNKMTVP-----KXKSTDESSGSDGNSFYTOMGNETSNGRG 358  
 Db 434 RELITAK--LQKPMSEKLFYSLLQYKQHSISLSSSENKKSAT----- 478  
 QY 359 RVLODAEERPHSRYLRAVSSDSGTSNSQSAKTYTMERCHAEMLSVSKSG-GENE 417  
 Db 479 -----ESSVNEPRITYASK--TANNITGLRSENNVYKTHLSLEIHEDSTYVONNALTGVNT 533  
 QY 418 ERSPDNNANINFEKERTSSSGSFERPDNNQALSNHL-CPGKPPFPAD----- 468  
 Db 534 EINAPV-----LAQKQFISINTLSQESDKAABAVTLPLPALPINASSRIFRNS 584  
 QY 469 -----PTPQETVOQWGNQIINAHAKTEHYDISINRPOGHPIQKDTKSNAMTDKV 524  
 Db 585 YTSISRSRSLRLSLNSRLSLSTSRKETHDNEM-----LPOLPKSPHRSYLSRRAT 638  
 QY 525 KNSDASDNASHVQKONTMYKMTALSHKEEIIQOECVFGSDPLSBSQSKRGMEPPWQGN 584  
 Db 639 HNSPSTKSHKISIRKNIATVAAARTLQNSAKSLVLSQISIKRS-----LNL 688  
 QY 585 RTLRSTPLVAHRLKPIQKTKKAVNSILDG-----EEVCVELV----- 624  
 Db 689 NDLVFPDPLPSK--KPASENVNKPSEHSLSDSDFEILCDQILFGNALDRILEEEDNE 746  
 QY 625 KEVASQEVYKVLQIISDGNIT--TIYPNGRGFPPLA-----DRPSPTDINSRYSF 675  
 Db 747 KERDTRQORQNDTKSSADFTISGVTKENEGPEPYFKIEKQPNMSYKSENNSGISS 806  
 QY 676 DNLPEKRYAKYASRFQVLYVSKSPKTYTFRYAKTIMENSPGADFEVWFYGVKLIHK 735  
 Db 807 PFIPEK-----ENTLSSSYEEOKPKRAALSDITNSFNKNNQ-----EGMRLEK 851  
 QY 736 TEDFIOVLEKTSYTLK-----SESEVNSLKEE--IGWMD-HANEGRICLALESITIS 787  
 Db 852 KIORQLOKXNDREPLKPIQHOELKRVSLPNDQKESLSDPRRNSIQPNNSYVESILQ 911  
 QY 788 EBERTRBAPPPPIIG-----RKPGSTSPKALSPPEVDSNY-PTDRASFN 836  
 Db 912 GKLFKEKPEASHWHERGSLFMSHEVEDKPYKASD-----VSISSYPLDTVATSSR 964  
 QY 837 MYMHAASPT-OAPI-----LNPSTWNGELGLTTASGTD----- 871  
 Db 965 DSVTLASSTIQKWLSPSSFLNTSM-TFNKLSQILLADGDDHGLSVPOKQSVAMSH 1023  
 QY 872 -TSSNSLKDCLPKSAQLKSVFK-----NYGMATOL-----TSGAV 907  
 Db 1024 PLRKQSAKISILTPRSNINANLNVKRGQSGPSTYLSNDLDSITMTPTAMEIPTTTAQT 1083

QY 908 WVOFND-----GSQLVQAGVSSISYTS-----PNCQTTYGENEKLDPY 947  
 Db 1084 QLMANDTNNKNTSPKASFTKEKVTLSAAITISKEKEPPNSDTNY-----IPDY 1133

RESULT 14  
 ID PILOI SCHPO STANDARD; PRT; 683 AA.  
 AC P50528;  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 GN PILOI OR SPAC23C11.16.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetes; Schizosaccharomycetaceae;  
 OC NCB1\_TaxID=4896;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=95262899; PubMed=7744248;  
 RX Okura H., Hagan I.M., Glover D.M.;  
 RT "The conserved Schizosaccharomyces pombe kinase p101, required to  
 form a bipolar spindle, the actin ring, and septum, can drive septum  
 formation in G1 and G2 cells."  
 RL Genes Dev. 9:1059-1073 (1995).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RX Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 Sgourou G., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
 Gentles S., Goble A., Hamlin N., Harris P., Hidalgo J., Hoagson G.,  
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,  
 Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,  
 Skelton J., Simmonds M., Squares D., Squares S., Stevens K.,  
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 Woodward J., Voicakert G., Aert R., Robben J., Grynolprez B.,  
 Meltjens I., Vanstreels E., Rieger M., Schaefer M., Meiller-Auer S.,  
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 Eger P., Zimmermann W., Wedler H., Reinhardt R., Pohl T.M.,  
 Goffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Motter S.,  
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
 Baga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,  
 Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potaashkin J.,  
 Spharakaki G.V., Usey D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880 (2002).  
 CC -!- FUNCTION: Required to form a bipolar spindle, the actin ring and  
 septum. Functions upstream of the whole septum formation pathway,  
 including actin ring formation (regulated by late septation genes)  
 and septal material deposition (regulated by early septation  
 genes). Behaves as a "septum-promoting factor", and could also be  
 involved in inducing other late events of cell division.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDCs/Polo subfamily.  
 CC -!- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85758; CAA59766.1; -.
DR EMBL; Z98559; CAB1167.1; -.
DR PIR; T38254; T38254.
DR HSBP; Q63450; 1A06.
DR Genedb SPombe; SPAC23C11.16; -.
DR InterPro; IPR000959; POLO_box.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00659; POLO_box; 2.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50078; POLO_Box; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat.
FT DOMAIN 41 296 PROTEIN_KINASE.
FT NP_BIND 47 55 ATP (BY SIMILARITY).
FT BINDING 69 69 ATP (BY SIMILARITY).
FT ACT_SITE 163 163 BY SIMILARITY.
FT DOMAIN 500 567 POLO_BOX 1.
FT DOMAIN 604 670 POLO_BOX 2.
FT SEQUENCE 683 AA; 77301 MW; F1ICDDEF9B913917 CRC64;

```

Query Match	10.1%	Score 514.5;	DB 1;	Length 683;
Best Local Similarity	24.4%	Pred. No. 3.8e-20;		
Matches 176;	Conservative 106;	Mismatches 247;	Indels 191;	Gaps 22;

```

Qy 18 GKSPGAVNPAEASHHGLVEALITIMDKKAYKAGMVOQVONEXKHQOLGHPISILEYN 77
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 47 IGBGGFARCFPVKQ- NYGNITYAAKVIARISLQNDKTKLKGELIGEHOSHMSHPINVEGDI 105
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 78 YFEDSNVYVLEVMCHNGENMRVYLNRYKPEPSENEARHPHQIITYGMULYHSHGILHRDL 137
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 106 CFEBSJNIVYLBELCEHKSIMELLRRK- KQTEPEVRIAMQIIGALKVMHKKGVIIHRDL 164
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 138 TLSVLLTRNNKNTKIADFGLATOLKMPHEKAYITLQGPNTYISPEIA- TRSAHGLESDVW 195
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 165 KLGIMIMDESNNVYIGDFGLAALLMDDEBRKMTIGCPNTYIAPBILFNFSKSGHFEVDLW 224
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 196 SLGCMFTYILGRPPFDTDVYKNTLKNVLDYEMPS- FLSIEKDLIHQLLRRNPARD 253
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 225 SAGVYVATALLGKPPQDKEVYKTYRKIKANSYSPSPSVNDISAEKDLISLLYHHD8IR 284
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 254 LSLSSVLDHPMWSRNSSTKSLQGLTVEBDSIDSGHATISTAITASSSTISGSLPDKRL 313
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 285 PSIDIDIYDHEF-----FHTGYMAST----- 304
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 314 IGOGLPKMNTVPRKXKST-----DPSGSDGNSFTYQW-----NQETS- 354
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 305 LPDEILHSMPIWPSQSQSKSFQRMULDPAASAGVGFGNSADEVKOKPYALFTDEVDNDRI 364
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 355 -----SGRGRIYQ-----DAERPHSRVLYRAYSSDRGSTNS-----QSGA- 391
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 365 LPSVLSPRDRNVPWIKIGPEFTKVPYKSLKSTLHARKSTODSLSRVKVLREBSQSFPVT 424
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 392 KTYMERCHSAEMLSVSKSGGGENEERYSTPDNNANFNFPKEXKTSSSSSSFERPDNQ 451
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 425 KSAVTEOUEPIQLI-----RSLSANVTSLSKVGMKSDIWSYKTKAKIG----- 471
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 452 ALSNHLCPGKTPFPADPTPQTEFVQWFGNQLINAHRLKTEYDSIS----- 499
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 472 -----MALEAHTHALTSBDADEBPVLFTKWVDY 500
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 500 PNRDFOGHPDLQKDTSKNAWTDTKVKNDSADNHSVKOONT-----MKYTAHL 550
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

```

D501 SNRYGAGYQLDSGVHFNDDTSLTFADAEVVEYALHPDOTEIKPIYIPASVSPESIR 560

D551 SKREIIQQGCVSGSPDLEEQSTRGMEPMGONRLLTSITSPVLARLRPIROKT---- 606

D561 SKQLLHKRXYSTMGNLSRAVDSEFKP---KST--SNTMLEMGIYLR--TRQALMFR 614

D607 -----KAVVS-----IDSEVCYLVKAYSQ-----EYKEVLQ 638

D615 SNGIFQFNFLDRKVVISTARKIIVLDEKEERVELPQEQSNAPSEDLSRLKRTITLE 674

RESULT 15

ID	STK6_HUMAN	STANDARD:	PRT: 403 AA.
AC	014965; 060445; 075873; 09B06; 09UP65;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Serine/threonine kinase 6 (EC 2.7.1.37) (Serine/threonine kinase 15) (AURA)		
DE	(Aurora/Aurp1-related kinase 1) (Aurora-related kinase 1) (AARK1)		
DE	(Aurora/A) (Breast-tumor-amplified kinase).		
GN	STK6 OR STK15 OR AIK OR ARK1 OR AURA OR BTK.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Blood;		
RC	MDLINE=97298093; PubMed=9153211;		
RA	Kikawa M., Kotani S., Hattori T., Sumi N., Yoshioke T., Todokoro K., Okano Y.;		
RT	"cell cycle-dependent expression and spindle pole localization of a novel human protein kinase, Aik, related to Aurora of Drosophila and yeast Ipl1.";		
RT	J. Biol. Chem. 272:13766-13771(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	MDLINE=98183439; PubMed=9514916;		
RA	Shindo M., Nakano H., Kuroyanagi H., Shiraawa T., Mihara M., Gilbey D.V., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;		
RT	"DNA cloning, expression, subcellular localization, and chromosomal assignment of mammalian aurora homologues, aurora-related kinase (ARK) 1 and 2.";		
RT	Biochem. Biophys. Res. Commun. 244:285-292(1998).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Breast;		
RC	MDLINE=98442657; PubMed=971714;		
RA	Zhou H., Kuang J., Zhong L., Kuo W.-L., Gray J.W., Sahin A., Brinley B. R., Sen S.;		
RT	"Tumour amplified kinase STK15/BTK induces centrosome amplification, aneuploidy and transformation.";		
RT	Nat. Genet. 20:189-193(1998).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Wang L., Thibodeau S.N.;		
RT	"functional analysis of the STK15 gene in human tumors.";		
RT	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	MDLINE=21638749; PubMed=11780052;		
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavridis K., Almeida J.P., Babbage A.K., Bagge C.L., Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.B., Bridgman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,		

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.  
RA Kay M.P., Kimberley A.M., King A., Knighes A., Laird G.K., Lawlor S.,  
RA Lehrsvalaino M.H., Leverisha M., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McComachie I.J., Mcley K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin T.C., McKernon T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,  
RA Swann R.M., Symcote N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilmington J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J. ;  
RT "The DNA sequence and comparative analysis of human chromosome 20." ;  
RL Nature 414:665-671(2001).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cervix, Colon, Kidney, and Muscle;  
RX MEDLINE=22389257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Altshuler S.P., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschuld R.P., Zeeberg B., Bueltow K.H., Scheefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hatoh F.,  
RA Datchenko L., Matrasina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein W.J., Ussid T.B., Toshynkyi S., Carlini P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gutarane P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,  
RA Rahay J., Heitlon B., Ketteman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Radermacher A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schmechel A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences." ;  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RP CELL-CYCLE REGULATION.  
RX MEDLINE=21895866; PubMed=11790771;  
RA Tanaka M., Ueda A., Kanamori H., Ideguchi H., Yang J., Kitajima S.,  
RA Ishiguro Y.;  
RT "Cell-cycle-dependent regulation of human aurora A transcription is  
mediated by periodic repression of E2F1." ;  
RL J. Biol. Chem. 277:10719-10726(2002).  
RN [8]  
RP REVIEW.  
RX MEDLINE=21306577; PubMed=11413462;  
RA Niig E.A.;  
RT "Mitotic kinases as regulators of cell division and its checkpoints." ;  
RL Nat. Rev. Mol. Cell Biol. 2:21-32(2001).  
CC -I- FUNCTION: May play a role in cell cycle regulation during anaphase  
and/or telophase, in relation to the function of the  
centrosome/spindle pole region during chromosome segregation.  
May be involved in microtubule formation and/or stabilization. May  
play a key role during tumor development and progression.  
CC -I- CARLYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -I- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells  
and at each spindle pole in mitosis.  
CC -I- TISSUE SPECIFICITY: Highly expressed in testis and weakly in  
colon, ovarian, prostate, neuroblastoma, breast and cervical  
cancer cell lines. Expression is cell-cycle regulated, low in  
G1/S, accumulates during G2/M, and decreases rapidly after.  
CC -I- PTM: Phosphorylated.  
CC -I- DISEAS: Defects in STK6 are responsible for numerical centrosome  
aberrations including aneuploidy.  
CC -I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC Aurora subfamily.  
CC -I- CAUTION: Ref.1 sequence differs from that shown due to frameshifts

Db	249	SKRVIRHDIKPELILGSAGELKIADFGWS--VHAPSSRRITLGGTLDYLPPEMIEGRWH	306
Qy	189	GUESDVWSIGCMFYTLIGRPFDITVKNLKVVLADYEMPSFISIAKDIHQULRR	248
Db	307	DERVDLWSLGVLCYEFLVSKPPEBANTYQETVKRISRVEFTPPDPVTEGARDLISRILKH	366
Qy	249	NPADRLSLSSVLDHPFMSRNSSTKS	273
Db	367	NPSQRPMLKREVLHPWITANSKPS	391

Search completed: September 23, 2004, 21:00:51  
Job time : 18 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:58:06 / Search time 63 Seconds  
(without alignments)  
4857.979 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078  
Sequence: 1 MATCHEDPKVGNLGLK.....KLQCLSLILMSNPTRPH 970

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_ricent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5074	99.9	970	4	Q81YF0
2	5067	99.8	970	4	Q00444
3	5057	99.6	970	4	Q96095
4	3927.5	77.3	925	11	Q64702
5	3918.5	77.2	925	11	Q8R015
6	2116	41.7	535	11	Q9CVU6
7	1583.5	31.2	324	11	Q9CVR6
8	1408.5	26.4	769	5	Q97143
9	1341.5	26.4	769	5	Q97143
10	1338.5	26.4	769	5	Q86N18
11	610	12.0	160	11	Q86N18
12	593.5	11.7	978	5	Q86N17
13	582	11.7	663	13	Q7X26
14	590.5	11.6	646	4	Q96CV1
15	587.5	11.6	504	11	Q8R017
16	579.5	11.4	644	13	Q7ZVM4

17	572.5	11.3	666	13	Q90XS4	Q90XS4 xenopus lae
18	571	11.2	557	13	Q90XS3	Q90XS3 xenopus lae
19	563.5	11.1	682	11	Q8K226	Q8K226 mus musculu
20	560.5	11.0	372	4	Q8R7M6	Q8R7M6 homo sapien
21	560	11.0	767	5	Q15882	Q15882 trypanosoma
22	543.5	10.7	623	5	Q81U35	Q81U35 aseterina pe
23	542.5	10.7	1075	5	Q95U75	Q95U75 drosophila
24	539	10.6	526	6	Q9BDP8	Q9BDP8 sus scrofa
25	538	10.6	1267	3	Q7Z865	Q7Z865 kluveromyc
26	537.5	10.6	502	10	Q8LR00	Q8LR00 oryza sativ
27	529	10.4	488	5	Q8SM6	Q8SM6 encephalico
28	527.5	10.4	598	13	Q7ZM06	Q7ZM06 xenopus lae
29	527	10.4	461	10	Q82051	Q82051 sorghum bic
30	527	10.4	1192	5	Q9TW45	Q9TW45 caenorhabd
31	527	10.4	1192	5	Q17346	Q17346 caenorhabd
32	525.5	10.3	598	13	P70032	P70032 xenopus lae
33	524.5	10.3	1033	5	Q8MLJ7	Q8MLJ7 drosophila
34	523.5	10.3	833	5	Q8SSX5	Q8SSX5 dictyostell
35	518	10.2	1066	5	Q8MVX0	Q8MVX0 haemonchus
36	517	10.2	565	5	Q8MLJ6	Q8MLJ6 drosophila
37	517	10.2	942	5	Q8MVM9	Q8MVM9 haemonchus
38	516.5	10.2	832	5	Q963B6	Q963B6 drosophila
39	516.5	10.2	1060	5	Q9V8W0	Q9V8W0 drosophila
40	515.5	10.2	582	5	Q9GRB7	Q9GRB7 hemiceurot
41	513	10.1	706	5	Q9V8V9	Q9V8V9 drosophila
42	513	10.1	938	5	Q9V8V8	Q9V8V8 drosophila
43	512	10.1	1046	5	Q7YU80	Q7YU80 drosophila
44	511.5	10.1	914	5	Q19469	Q19469 caenorhabd
45	508.5	10.0	1058	5	Q963B5	Q963B5 drosophila

## ALIGNMENTS

RESULT 1

Q81YF0 PRELIMINARY; PRT; 970 AA.

AC Q81YF0; 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Similar to serine/threonine protein kinase.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC036023; ANH36023.1; -

DR GO; GO:0005524; P:ATP binding; IEA.

DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.

DR GO; GO:0007049; P:cell cycle; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000959; POLO box.

DR InterPro; IPR000719; Prot\_kinase.

DR InterPro; IPR002290; Ser\_thr\_kinase.

DR InterPro; IPR001245; Tyr\_kinase.

DR InterPro; IPR008266; Tyr\_kinase\_AS.

DR Pfam; PF00069; pkinase; 1.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00220; S\_TKc; 1.

DR SMART; SM00219; TYKc; 1.

DR PROSITE; PSS0078; POLO\_BOX; 1.

DR PROSITE; PSS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PSS00109; PROTEIN\_KINASE\_TYR; 1.

DR Kinase; Serine/threonine-protein kinase.

SEQUENCE 970 AA; 108971 MW; 4D56F5FD983211A6 CRC64;

Query Match 99.9%; Score 5074; DB 4; Length 970;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60
DB 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60
QY 61 VKIHQQLKHPSTILELYNFEDSNVYVLVLEMGCHNGEMNRYLKNRYKPSSEBARHFMHQI 120
DB 61 VKIHQQLKHPSTILELYNFEDSNVYVLVLEMGCHNGEMNRYLKNRYKPSSEBARHFMHQI 120
QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180
DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180
QY 181 EIATRSAGLESVDWSLGCMEFTYLLIGRPDPDTIVKNTLNKVLADYEMPSFLSTEAKD 240
DB 181 EIATRSAGLESVDWSLGCMEFTYLLIGRPDPDTIVKNTLNKVLADYEMPSFLSTEAKD 240
QY 241 LIHOLLRRNPADRLSLSSVLDPHFMRSNSTSKDGLTVEDSIDSGHATITATITASSST 300
DB 241 LIHOLLRRNPADRLSLSSVLDPHFMRSNSTSKDGLTVEDSIDSGHATITATITASSST 300
QY 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGNFYTOMGNOETNSNGRGV 360
DB 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGNFYTOMGNOETNSNGRGV 360
QY 361 IODAEERPHSRYLRRAYSDDSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENERY 420
DB 361 IODAEERPHSRYLRRAYSDDSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENERY 420
QY 421 SPTDNNAINFNFPEKETSFGSFERPDNNQALSNHLCPGKTPPPADPTPOTETVQOQF 480
DB 421 SPTDNNAINFNFPEKETSFGSFERPDNNQALSNHLCPGKTPPPADPTPOTETVQOQF 480
QY 481 GNLQINAHLRKTEYDLSIPNRDFOQHPLQKDTSKNMTDTKVKKNSDASDNASHVKQO 540
DB 481 GNLQINAHLRKTEYDLSIPNRDFOQHPLQKDTSKNMTDTKVKKNSDASDNASHVKQO 540
QY 541 NNMKWTALHSHKPEIIOQECVFGSDPLSROSKTRGMEPWGQONTLRISITPLVAHRLK 600
DB 541 NNMKWTALHSHKPEIIOQECVFGSDPLSROSKTRGMEPWGQONTLRISITPLVAHRLK 600
QY 601 PIROKTKAVVAILDSEEVCELVKESYASOEYKEVLQISSDQNTTITYPNGRGFPPLA 660
DB 601 PIROKTKAVVAILDSEEVCELVKESYASOEYKEVLQISSDQNTTITYPNGRGFPPLA 660
QY 661 DRPSPPTDNISRYSPDNLPKTYWKRYQVARSFVQVLRSHKSPKITTYFTRAYAKCILMENSFG 720
DB 661 DRPSPPTDNISRYSPDNLPKTYWKRYQVARSFVQVLRSHKSPKITTYFTRAYAKCILMENSFG 720
QY 721 ADPEWFWYGVKVIHKTEDFIQVIEKTKGKSYTLKSBEVNSLKEELIKMYVDHANEGHRLCL 780
DB 721 ADPEWFWYGVKVIHKTEDFIQVIEKTKGKSYTLKSBEVNSLKEELIKMYVDHANEGHRLCL 780
QY 781 ALESTIISSEERKTRSAFPFPIIIGRKPSTSPKALSPPSVDNSNYPTDRASFNRMVWH 840
DB 781 ALESTIISSEERKTRSAFPFPIIIGRKPSTSPKALSPPSVDNSNYPTDRASFNRMVWH 840
QY 841 SAASPTQADILNPSVNTNEGLGLTTTASGTDIISNSLKCPLPKSAQLLKSVEFKNVGMAT 900
DB 841 SAASPTQADILNPSVNTNEGLGLTTTASGTDIISNSLKCPLPKSAQLLKSVEFKNVGMAT 900
QY 901 QLTSAVWVQFNDGSLVYQAGVSSISYTPSPNGQTRVGENEKLDPYIKQKQCLSSITLL 960
DB 901 QLTSAVWVQFNDGSLVYQAGVSSISYTPSPNGQTRVGENEKLDPYIKQKQCLSSITLL 960
QY 961 MFSNPTPNFH 970
DB 961 MFSNPTPNFH 970

```

## RESULT 2

```

ID 000444 PRELIMINARY; PRT; 970 AA.
AC 000444;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase.
GN SAK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Karn T., Holtrich U., Wolf G., Hock B., Streibhardt K.,
RA Ruebsamen-Waigmann H.;
RT "Human SAK related to the PLK/polo family of cell cycle kinases shows
RT high mRNA expression in testis.";
RL Oncol. Rep. 4:505-510(1997).
DR EMBL; Y13115; CAA73575.1; -.
DR HSSP; O63450; 1A06.
DR Genew; HGNC:11397; STK18.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000559; POLO_box.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase.1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50078; POLO_BOX; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase.
SQ
SEQUENCE 970 Aa; 10895 Mw; 552535615FDBB8DF CRC64;

```

Query Match 99.8%; Score 5067; DB 4; Length 970;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 967; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60
DB 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60
QY 61 VKIHQQLKHPSTILELYNFEDSNVYVLVLEMGCHNGEMNRYLKNRYKPSSEBARHFMHQI 120
DB 61 VKIHQQLKHPSTILELYNFEDSNVYVLVLEMGCHNGEMNRYLKNRYKPSSEBARHFMHQI 120
QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180
DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180
QY 181 EIATRSAGLESVDWSLGCMEFTYLLIGRPDPDTIVKNTLNKVLADYEMPSFLSTEAKD 240
DB 181 EIATRSAGLESVDWSLGCMEFTYLLIGRPDPDTIVKNTLNKVLADYEMPSFLSTEAKD 240
QY 241 LIHOLLRRNPADRLSLSSVLDPHFMRSNSTSKDGLTVEDSIDSGHATITATITASSST 300
DB 241 LIHOLLRRNPADRLSLSSVLDPHFMRSNSTSKDGLTVEDSIDSGHATITATITASSST 300
QY 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGNFYTOMGNOETNSNGRGV 360
DB 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGNFYTOMGNOETNSNGRGV 360
QY 361 IODAEERPHSRYLRRAYSDDSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENERY 420
DB 361 IODAEERPHSRYLRRAYSDDSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENERY 420

```

```

Db 361 IODAEERPHSRYLRLAYSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
Qy 421 SPTDNNANIFNFPEKETTSSSGSFERRPDNNQALSNHLCGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNFPEKETTSSSGSFERRPDNNQALSNHLCGKTPFPADPTPOTETVQOMF 480
Qy 481 GNLIQINAHARKTTEYDSISPNRDFOGHPDLQKDTSGNAATDTVKKGSADSDAHSVKQO 540
Db 481 GNLIQINAHARKTTEYDSISPNRDFOGHPDLQKDTSGNAATDTVKKGSADSDAHSVKQO 540
Qy 541 NTMKWTALHSKPEIIQOECVFGSDPLSEOSKTRGMEPPMGVONRLRSITSPVLVHRLK 600
Db 541 NTMKWTALHSKPEIIQOECVFGSDPLSEOSKTRGMEPPMGVONRLRSITSPVLVHRLK 600
Qy 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITTIYYPNNGRGFPPLA 660
Db 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITTIYYPNNGRGFPPLA 660
Qy 661 DRPPSPTDNISRYSPNLPPEKTRWKYQVARSFVQVLRSPKITYPTRAKCLIMENSPG 720
Db 661 DRPPSPTDNISRYSPNLPPEKTRWKYQVARSFVQVLRSPKITYPTRAKCLIMENSPG 720
Qy 721 ADPEWVFYDGVKIHKTEDFIOVIEKTKGKYTLKSEBVSNSLKEEIKMYMDHANEGHRI 780
Db 721 ADPEWVFYDGVKIHKTEDFIOVIEKTKGKYTLKSEBVSNSLKEEIKMYMDHANEGHRI 780
Qy 781 ALESIISEERKTRSAFPFPIIIGKRGSTSSPKALSPPPSVDSNPTDRAFSFNMVNM 840
Db 781 ALESIISEERKTRSAFPFPIIIGKRGSTSSPKALSPPPSVDSNPTDRAFSFNMVNM 840
Qy 841 SAASPTQALINPSWMTNGLGLTTASGTDISSNLKQCLPKSAQLKLSVFKNVGMAT 900
Db 841 SAASPTQALINPSWMTNGLGLTTASGTDISSNLKQCLPKSAQLKLSVFKNVGMAT 900
Qy 901 QLTSGAVWQFNDGSQLVVOAGVSSISYTSPPNGQTRYGENSEKLPYIYIKQKQCLSSILL 960
Db 901 QLTSGAVWQFNDGSQLVVOAGVSSISYTSPPNGQTRYGENSEKLPYIYIKQKQCLSSILL 960
Qy 961 MFSNPTNFH 970
Db 961 MFSNPTNFH 970

```

```

RESULT 3
Q96Q95 PRELIMINARY; PRT; 970 AA.
ID Q96Q95;
AC Q96Q95;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Sak.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=21486437; PubMed=11489907;
RA Yamaehita Y., Kajigaya S., Yoshida K., Ueno S., Ota J., Ohmine K.,
RA Ueda M., Miyazato A., Ohya K., Kitamura T., Ozawa K., Mano H.;
RT "Sak Serine-Threonine Kinase Acts as an Effector of Tec Tyrosine
RT Kinase.";
RT J. Biol. Chem. 276:39012-39020 (2001).
DR HSSP; P24941; 1BUH.
DR GO:GO:0005524; P:ATP binding; IEA.
DR GO:GO:0004774; P:protein binding/threonine kinase activity; IEA.
DR GO:GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO:GO:0016740; P:transferase activity; IEA.
DR GO:GO:0007049; P:cell cycle; IEA.
DR GO:GO:0006688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000959; POLO_box.

```

```

DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser thr kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00078; POLO_BOX; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Transferase.
SQ SEQUENCE 970 AA; 109084 MW; 1F431PA607A3550A CRC64;

Query Match 99.6%; Score 5057; DB 4; Length 970;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 965; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MATCI GEKIEDPKVGNLKGKSPAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQVRQNE 60
Db 1 MATCI GEKIEDPKVGNLKGKSPAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQVRQNE 60
Qy 61 VKIHQQLKHPSTILELYNFEEDSNYYVLVLEMCHGEMNRYLKNRVKPFSENEARHFMQI 120
Db 61 VKIHQQLKHPSTILELYNFEEDSNYYVLVLEMCHGEMNRYLKNRVKPFSENEARHFMQI 120
Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKADRLATQLKCPHEKHYTLGCTPYVISP 180
Db 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKADRLATQLKCPHEKHYTLGCTPYVISP 180
Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKADRLATQLKCPHEKHYTLGCTPYVISP 180
Db 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKADRLATQLKCPHEKHYTLGCTPYVISP 180
Qy 181 EIAATSAHGLSDVWSIGCMFTLLIGRPPPDVYKNTLKNVVLADYEMPSFLSTEAKD 240
Db 181 EIAATSAHGLSDVWSIGCMFTLLIGRPPPDVYKNTLKNVVLADYEMPSFLSTEAKD 240
Qy 241 LIHQLLRNPADRLSLSSVLDHPFMSRNSSTKSLGTVEDSIDSGHATISTAITASSST 300
Db 241 LIHQLLRNPADRLSLSSVLDHPFMSRNSSTKSLGTVEDSIDSGHATISTAITASSST 300
Qy 301 S1SGSLFDKRLILIGPLPNKNTVFPKAKSSITDFSSSGGNSFTYWGQOEFSNSGRGV 360
Db 301 S1SGSLFDKRLILIGPLPNKNTVFPKAKSSITDFSSSGGNSFTYWGQOEFSNSGRGV 360
Qy 361 IODAEERPHSRYLRLAYSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
Db 361 IODAEERPHSRYLRLAYSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
Qy 421 SPTDNNANIFNFPEKETTSSSGSFERRPDNNQALSNHLCGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNFPEKETTSSSGSFERRPDNNQALSNHLCGKTPFPADPTPOTETVQOMF 480
Qy 481 GNLIQINAHARKTTEYDSISPNRDFOGHPDLQKDTSGNAATDTVKKGSADSDAHSVKQO 540
Db 481 GNLIQINAHARKTTEYDSISPNRDFOGHPDLQKDTSGNAATDTVKKGSADSDAHSVKQO 540
Qy 541 NTMKWTALHSKPEIIQOECVFGSDPLSEOSKTRGMEPPMGVONRLRSITSPVLVHRLK 600
Db 541 NTMKWTALHSKPEIIQOECVFGSDPLSEOSKTRGMEPPMGVONRLRSITSPVLVHRLK 600
Qy 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITTIYYPNNGRGFPPLA 660
Db 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITTIYYPNNGRGFPPLA 660
Qy 661 DRPPSPTDNISRYSPNLPPEKTRWKYQVARSFVQVLRSPKITYPTRAKCLIMENSPG 720
Db 661 DRPPSPTDNISRYSPNLPPEKTRWKYQVARSFVQVLRSPKITYPTRAKCLIMENSPG 720
Qy 721 ADPEWVFYDGVKIHKTEDFIOVIEKTKGKYTLKSEBVSNSLKEEIKMYMDHANEGHRI 780
Db 721 ADPEWVFYDGVKIHKTEDFIOVIEKTKGKYTLKSEBVSNSLKEEIKMYMDHANEGHRI 780
Qy 781 ALESIISEERKTRSAFPFPIIIGKRGSTSSPKALSPPPSVDSNPTDRAFSFNMVNM 840
Db 781 ALESIISEERKTRSAFPFPIIIGKRGSTSSPKALSPPPSVDSNPTDRAFSFNMVNM 840

```

Db 781 ALBSIISEERKTSAPFFPIIIGRKPGSTSSPKALSPPSVDSNPTDRDASFNRMVMH 840  
QY 841 SAASPTQAPILNPMWNTNEGILTTTASGTDISSNSLKDCIPKSAQILKSVFNVMAT 900  
Db 841 SDASPTQAPILNPMWNTNEGILTTTASGTDISSNSLKDCIPKSAQILKSVFNVMAT 900  
QY 901 QLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTGENEKLPDYIKOKLOCLSSII 960  
Db 901 QLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTGENEKLPDYIKOKLOCLSSII 960  
QY 961 MFSNPTNPF 970  
Db 961 MFSNPTNPF 970  
RESULT 4  
064702 PRELIMINARY; PRT; 925 AA.  
ID 064702; AC 064702; DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DE 01-NOV-2003 (Tremblrel. 25, Last annotation update)  
GN SMK/PLK-AKIN kinase (protein kinase SMK/PLK-AKIN) (EC 2.7.1.1-).  
OS Mus musculus (mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DRA/2; TISSUE=Lymphoma;  
RX MEDLINE=94294387; PubMed=8022793;  
RA Fode C., Motro B., Yousefi S., Heffernan M., Dennis J.W.;  
RT "Sak, a murine protein-serine/threonine kinase that is related to the  
RT Drosophila polo kinase and involved in cell proliferation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:6388-6392(1994).  
CC -1- ALTERNATIVE PRODUCTS;  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Sak-a;  
CC IsoId=Q64702-1; Sequence=Displayed;  
CC Name=Sak-b;  
CC IsoId=Q64702-2; Sequence=VSP\_050448, VSP\_050449;  
CC EMBL: J29479; AAC37648.1; -;  
CC EMBL: L29480; AAC37649.1; -;  
CC PIR: A55748; A55748.  
DR HSBP: Q00534; IBI8.  
DR MGD; MGI:101783; Sck18.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0007049; F:cell cycle; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000959; POLO box.  
DR InterPro; IPR000719; POLO kinase.  
DR InterPro; IPR002290; Ser Thr kinase.  
DR InterPro; IPR008266; Ty\_r\_kinase\_AS.  
DR Pfam; PF00069; kinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; POLO BOX; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00101; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
KM ATP-binding; Alternative splicing; Serine/threonine-protein kinase;  
KM transferase.  
FT TRANSPLIC 417 464  
FT SNHCLGKTPPEPPADOTPEMVOQMGFNLOMNAHIGETN  
FT EHHVSP -> RYSPTEKSNVUTLSINTKQPIYVDLKKRI  
FT MTEQKDLNLTINKFDR (in isoform Sak-b).  
FT VARSPLIC 465 925  
FT Missing (in isoform Sak-b).  
FT Missing (in isoform Sak-b).  
FT Missing (in isoform Sak-b).  
FT Missing (in isoform Sak-b).  
SQ SEQUENCE 925 AA; 103685 MW; D868A76BB7343EB1 CRC64;

Query Match 77.3%; Score 3927.5; DB 11; Length 925;  
Best Local Similarity 78.6%; Pred. No. 2.2e-248;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;  
QY 1 MATCGEKTEDEKFNGLKGSFAGYVRAESHTGLVLAIKRIDKKAAYKAGVORVONE 60  
Db 1 MACIGERLEDEKFNGLKGSFAGYVRAESHTGLVLAIKRIDKKAAYKAGVORVONE 60  
QY 61 VKIHQKLPSTLLELYNPFEDSNVYLVLEMCHGEMNRYLKRVKPSSEBARHFMHOI 120  
Db 61 VKIHQKLPSTLLELYNPFEDSNVYLVLEMCHGEMNRYLKRVKPSSEBARHFMHOI 120  
QY 121 ITGMVYHSHGILHDLTSLNLTFRNNYIKADFGALATOLRMPHKKYTLGTPNYTSP 180  
Db 121 ITGMVYHSHGILHDLTSLNLTFRNNYIKADFGALATOLRMPHKKYTLGTPNYTSP 180  
QY 181 E1ATSAGHLESDVSLGCMFYTLIGRPPTDTVKNLTNKKVVLADYEMPAFLSREXOD 240  
Db 181 E1ATSAGHLESDVSLGCMFYTLIGRPPTDTVKNLTNKKVVLADYEMPAFLSREXOD 240  
QY 241 LIHQLLRNPDRLSLSSVLDHPMSRNSSTKSXDLGTVEDSIDGHATISTATTAASST 300  
Db 241 LIHQLLRNPDRLSLSSVLDHPMSRNSSTKSXDLGTVEDSIDGHATISTATTAASST 300  
QY 301 SLSGSLPDKRLLIGQPLPKCTYPPKXKSTDSSGDSGSPYTGNGN--QETSNSGRG 358  
Db 301 SLSGSLD-RRLLVQPLPKKITVFOKNKSSD-SSGDSGSPCTGNGNPEQENANSRG 358  
QY 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSVRSGGGENEE 418  
Db 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSVRSGGGENEE 418  
QY 419 RYSPIDNNANIFNPFKEKTSSSGSEFEPDNNQALSNHLCPGKTPPPADTPQTEYVQ 478  
Db 419 RYSPIDNNANIFNPFKEKTSSSGSEFEPDNNQALSNHLCPGKTPPPADTPQTEYVQ 478  
QY 479 WFGNLIQINAHLRKTTVEYDSTSPNRPQGHPIQDTSKMAWTDKVKKNSDASNNASVX 538  
Db 479 WFGNLIQINAHLRKTTVEYDSTSPNRPQGHPIQDTSKMAWTDKVKKNSDASNNASVX 538  
QY 539 QONTMKTMTALSHKPIIIOECVFGSDPLSEOSKTRGMEPPMGYONTLTSITSPVAHR 598  
Db 539 QONTMKTMTALSHKPIIIOECVFGSDPLSEOSKTRGMEPPMGYONTLTSITSPVAHR 598  
QY 599 LKPIROKTKKAVVSIIDSEEVCELVKYEASQRYKAVLOISDQNTITTYPNGRGFP 658  
Db 599 LKPIROKTKKAVVSIIDSEEVCELVKYEASQRYKAVLOISDQNTITTYPNGRGFP 658  
QY 659 LADRPSPPTNISRYSFDNLPEKWRKYQYASRFVOLVRKSPKITTYFTFYACILMENS 718  
Db 659 LADRPSPPTNISRYSFDNLPEKWRKYQYASRFVOLVRKSPKITTYFTFYACILMENS 718  
QY 719 PGADFEVWFDGKVIKHTEDFIOVIEKTKGKSYTLKSESVNSLKEEIKMYMDHANEGR 778  
Db 719 PGADFEVWFDGKVIKHTEDFIOVIEKTKGKSYTLKSESVNSLKEEIKMYMDHANEGR 778  
QY 779 CLAESIISEERKTSAPFFPIIIGRKPGSTSSPKALSPPSVDSNPTDRDASFNRMVMH 838  
Db 779 CLAESIISEERKTSAPFFPIIIGRKPGSTSSPKALSPPSVDSNPTDRDASFNRMVMH 838  
QY 839 MBSAASPTQAPILNPMWNTNEGILTTTASGTDISSNSLKDCIPKSAQILKSVFNVMAT 898  
Db 839 MBSAASPTQAPILNPMWNTNEGILTTTASGTDISSNSLKDCIPKSAQILKSVFNVMAT 898  
QY 899 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTGENEKLPDYIKOKLOCLSSI 958  
Db 899 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTGENEKLPDYIKOKLOCLSSI 958  
QY 959 LMFNSPTNPF 969  
Db 959 LMFNSPTNPF 969  
QY 914 LMFNSPTNPF 924  
Db 914 LMFNSPTNPF 924

```

RESULT 5
Q8R015 PRELIMINARY; PRT; 925 AA.
ID Q8R015;
AC Q8R015;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN STK18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Struhsberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026785; AAH26785.1; -.
DR PDB; 1MBY; 30-OCT-02.
DR MGI; 101783; Sck18.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007049; P:cell cycle; IEA.
DR GO; GO:0004468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000959; POLO_box.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001230; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PSS0078; POLO_BOX; 1.
DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSS0109; PROTEIN_KINASE_TYR; 1.
DR Hypothetical protein; ATP-binding; Transferase.
SQ SEQUENCE 925 AA; 103861 MW; 844APFC9MAC54C1 CRC64;

Query Match 77.2%; Score 3918.5; DB 11; Length 925;
Best Local Similarity 78.4%; Pred. No. 8.6e-248;
Matches 761; Conservative 77; Mismatches 84; Indels 49; Gaps 9;

QY 1 MATCIQEKIEDPFVGNLKGKSPAGYRAISHTGLGVAKIMDKKAMYGAVQVRQNS 60
DB 1 MACICGERIEDKRVGNLKGKSPAGYRAISHTGLGVAKIMDKKAMYGAVQVRQNS 60
QY 61 VKIHQQLKHPISILELYNFEDSNYYVLVEMCHNGEMNRYLKNRVKFSSEARHFMQI 120
DB 61 VKIHQQLKHPISILELYNFEDSNYYVLVEMCHNGEMNRYLKNRVKFSSEARHFMQI 120
QY 61 VKIHQQLKHPISILELYNFEDSNYYVLVEMCHNGEMNRYLKNRVKFSSEARHFMQI 120
DB 61 VKIHQQLKHPISILELYNFEDSNYYVLVEMCHNGEMNRYLKNRVKFSSEARHFMQI 120
QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLAQLKPKHKHTYLCTPRYVISP 180
DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLAQLKPKHKHTYLCTPRYVISP 180
QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLAQLKPKHKHTYLCTPRYVISP 180
DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLAQLKPKHKHTYLCTPRYVISP 180
QY 181 EIAIRSAHGLIESDVMSLIGCMFYLLIGRPFDTDVYKNTLKVVLADYEMPSFLISLEAKD 240
DB 181 EIAIRSAHGLIESDVMSLIGCMFYLLIGRPFDTDVYKNTLKVVLADYEMPSFLISLEAKD 240
QY 181 EIAIRSAHGLIESDVMSLIGCMFYLLIGRPFDTDVYKNTLKVVLADYEMPSFLISLEAKD 240
DB 181 EIAIRSAHGLIESDVMSLIGCMFYLLIGRPFDTDVYKNTLKVVLADYEMPSFLISLEAKD 240
QY 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATITSAITASSST 300
DB 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATITSAITASSST 300
QY 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATITSAITASSST 300
DB 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATITSAITASSST 300
QY 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSSGGNSGYTQMGN--OETNSGRG 358
DB 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSSGGNSGYTQMGN--OETNSGRG 358
QY 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSSGGNSGYTQMGN--OETNSGRG 358
DB 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSSGGNSGYTQMGN--OETNSGRG 358
QY 359 RVYDAEERPHSRYLRAVSDRSGTNSQSAKTYTMRCHSAEMLSYKSRGGGENEB 418
DB 359 RVYDAEERPHSRYLRAVSDRSGTNSQSAKTYTMRCHSAEMLSYKSRGGGENEB 418

```

```

DB 359 RVYDAEERPHSRYLRAVSDRSGTNSQSAKTYTMRCHSAEMLSYKSRGGGENEB 418
QY 419 RYSPIDNNANINFPREKTSSSSGSERPDNNQALSNHLCPEKTPPEPADPTPOETVQO 478
DB 419 RYSPIDNNANINFPREKTSSSSGSERPDNNQALSNHLCPEKTPPEPADPTPOETVQO 478
QY 411 -----LDENQSSNNHCLCKTPEPPADQTPQEMEMVQO 442
DB 411 -----LDENQSSNNHCLCKTPEPPADQTPQEMEMVQO 442
QY 479 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNSDASDNASHYK 538
DB 479 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNSDASDNASHYK 538
QY 443 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNSDASDNASHYK 501
DB 443 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNSDASDNASHYK 501
QY 539 QONTMKTALHSKPEITIQECVFGSDPLSEQSKTRCMEPPMGYQNTLRSTSPVANH 598
DB 539 QONTMKTALHSKPEITIQECVFGSDPLSEQSKTRCMEPPMGYQNTLRSTSPVANH 598
QY 502 QLSAKMYGSAHHKKEVMPQEP--DLHPHSEOSKRSMESTLGYOKPTRLRSTSPILARH 559
DB 502 QLSAKMYGSAHHKKEVMPQEP--DLHPHSEOSKRSMESTLGYOKPTRLRSTSPILARH 559
QY 599 LKPIQKTKKAVVSLIDSEVCELVKEXASQEVYVQISDQGTITITYPNGRGPR 658
DB 599 LKPIQKTKKAVVSLIDSEVCELVKEXASQEVYVQISDQGTITITYPNGRGPR 658
QY 560 LKPIQKTKKAVVSLIDSEVCELVKEXASQEVYVQISDQGTITITYPNGRGPR 619
DB 560 LKPIQKTKKAVVSLIDSEVCELVKEXASQEVYVQISDQGTITITYPNGRGPR 619
QY 659 LADRPSPPTDNISRYSPFNLPEKYWKYQYASRFVQVYKSPKITYPTRYAKCILMENS 718
DB 659 LADRPSPPTDNISRYSPFNLPEKYWKYQYASRFVQVYKSPKITYPTRYAKCILMENS 718
QY 620 LADRPPLPTNNISRYSPFNLPEKYWKYQYASRFVQVYKSPKITYPTRYAKCILMENS 679
DB 620 LADRPPLPTNNISRYSPFNLPEKYWKYQYASRFVQVYKSPKITYPTRYAKCILMENS 679
QY 719 PGADPEWVYDGVKIHKTEDFQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 778
DB 719 PGADPEWVYDGVKIHKTEDFQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 778
QY 680 PGADPEWVYDGVKIHKTEDFQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 739
DB 680 PGADPEWVYDGVKIHKTEDFQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 739
QY 779 CLATESIIEEERKTRSAFPFLLTIGRKGRGSSPKALSPPSVSNYPTTRDASFNRY 838
DB 779 CLATESIIEEERKTRSAFPFLLTIGRKGRGSSPKALSPPSVSNYPTTRDASFNRY 838
QY 740 CLATESIIEEERKTRSAFPFLLTIGRKGRGSSPKALSPPSVSNYPTTRDASFNRY 798
DB 740 CLATESIIEEERKTRSAFPFLLTIGRKGRGSSPKALSPPSVSNYPTTRDASFNRY 798
QY 839 MESAASPTQAPILNPSMVTNEGILTTTASGTDISNSLKDCLPKSAQLLKSVFNVQVM 898
DB 839 MESAASPTQAPILNPSMVTNEGILTTTASGTDISNSLKDCLPKSAQLLKSVFNVQVM 898
QY 799 VNSAAPPFOSPGLSBETVVEGLGHTATATGCVSS-----LPSAQLLKSVFNVQVM 853
DB 799 VNSAAPPFOSPGLSBETVVEGLGHTATATGCVSS-----LPSAQLLKSVFNVQVM 853
QY 899 ATQLTSGAVVQFNDQSQVLVQAVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 958
DB 899 ATQLTSGAVVQFNDQSQVLVQAVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 958
QY 854 ATQLTSGAVVQFNDQSQVLVQAVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 913
DB 854 ATQLTSGAVVQFNDQSQVLVQAVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 913
QY 959 LMFSPNTPNF 969
DB 959 LMFSPNTPNF 924

RESULT 6
Q9CVU6 PRELIMINARY; PRT; 535 AA.
ID Q9CVU6;
AC Q9CVU6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Serine/threonine kinase 18 (Fragment).
GN STK18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=21085660; PubMed=11217651;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsumoto H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuenli P., Lewis S., Matsumoto Y., Nishikido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.U., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.",  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK006459; BAB24599.1; -  
 DR MGD; MGI:101783; Stk18.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007049; F:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; POLO box.  
 DR PROSITE; PSS0078; POLO\_Box; 1.  
 FT NON TER 1  
 SQ SEQUENCE 535 AA; 59788 MW; 3A35F0713FD641B3 CRC64;  
 Query Match 41.7%; Score 2116; DB 11; Length 535;  
 Best Local Similarity 71.6%; Pred. No. 4e-130;  
 Matches 414; Conservative 53; Mismatches 67; Indels 44; Gaps 5;  
 QY 392 KTYTMRCHSAEMLSVSKRGSGGSENERISPTNNANITNFFKKTSSSGSPRPDNNQ 451  
 DB 1 KTVSECHSVEMLSKPRRS-----LDENQ 25  
 QY 452 ALSNHLCPGKTPFPADPTPQTEVQVQFQNLQINHLKRTTEYDSISPNRDFQHPDQ 511  
 DB HSNHHCLGKTPFPADPTPQTEVQVQFQNLQINHLKRTTEYDSISPNRDFQHPDQ 85  
 QY 512 KQTSKNAATDTKTKKSDASDNASHYKQONTKMTALSKREIIQOECVFGSDPLSEOS 571  
 DB 86 -DILRNAMDTDRASKADTSANVAHVAKOLSAMKYVAHKKPEVMQBP--GLHPHSEGS 142  
 QY 572 KTEGMPWQYQRTLRSTSPVARRLRPIRQKTKKAVSLIDSEVCELVKYEASG 631  
 DB 143 KRSNEMSTGYQPTLRSTSPVARRLRPIRQKTKKAVSLIDSEVCELVKYEASG 202  
 QY 632 YKAEVLQISSDGTITTYYPNGRGFPLADRPSPPTDNISSYFQNLPEKRYKTYQASR 691  
 DB 203 YKAEVLQISSDGTITTYYPNGRGFPLADRPSPPTDNISSYFQNLPEKRYKTYQASR 262  
 QY 692 PQVLVSKSPKITYFRYAKKILIMENSPPADPEWVYDQVTKHTEDFQVTEKTSYT 751  
 DB 263 FQVLVSKSPKITYFRYAKKILIMENSPPADPEWVYDQVTKHTEDFQVTEKTSYT 322  
 QY 752 LKSESEVNSLKEIKVMYMDHANEGRICLSEIISSEERKTRSAFPPIIIGRRDGS 811  
 DB 323 LKNEVEVTLKEVEKVMYMDHANEGRICLSEIISSEERKTRSAFPPIIIGRRDGS 382  
 QY 812 SPKALSPSPSVUNSPTRDRASFNRVWMSAASPQADPLINSMVTNGLGLTTASGTD 871  
 DB 383 SPKALSPAP-VDPSPCCGEQASRLSVNSAAPTQSPQLSPSTVYVGLGHTATATATG 441  
 QY 872 ISSNSLKDCLPKSAOLKSFVYKVGATQNLGSAVWVQFNDGSLVQAGVSSISYSP 931  
 DB 442 VSSS-----LPSAOLKSFVYKVGATQNLGSAVWVQFNDGSLVQAGVSSISYSP 496  
 QY 932 NGQTRRYGENEKLDPYIKQKLOCLSSILLMFSNPTNF 969  
 DB 497 DGQTRRYGENEKLDPYIKQKLOCLSSILLMFSNPTNF 534  
 RESULT 7  
 Q9CVR6 PRELIMINARY; PRT; 324 AA.  
 AC Q9CVR6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Setine/threonine kinase 18 (Fragment).  
 GN STK18.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kanakawa T., Saito R.,  
 RA Kadocia K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gastinech S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli I., Sakamoto N.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.",  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK006827; BAB24759.1; -  
 DR HSP; Q00534; 1B18.  
 DR MGD; MGI:101783; Stk18.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002290; Ser Thr kinase.  
 DR InterPro; IPR001245; Tyr kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 1.  
 DR PRINTS; PR00101; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 FT ATP-binding; Transferase.  
 FT NON TER 324  
 SQ SEQUENCE 324 AA; 36568 MW; 04247065DFB198CC CRC64;  
 Query Match 31.2%; Score 1583.5; DB 11; Length 324;  
 Best Local Similarity 92.3%; Pred. No. 1.4e-95;  
 Matches 300; Conservative 15; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 MATCIKERIDFKVGNLKGKSPAGYRAESITGTEVAIKMIDKKAMYKAGVQVONE 60  
 DB 1 MAACIGRIDFDKVGNLKGKSPAGYRAESITGTEVAIKMIDKKAMYKAGVQVONE 60  
 QY 61 VKIHQQLKHSILIELNYFEDSNVYVLVLMCHNGEMNRYLKNRVYFSENEARHFMQI 120  
 DB 61 VKIHQQLKHSVLELNYFEDNNVYVLVLMCHNGEMNRYLKNRVYFSENEARHFMQI 120  
 QY 121 ITGMVLVHSHGILHARDLITSNLLITRNMIKTAADPLAQLQKMPHKHNTLCTPYIYP 180  
 DB 121 ITGMVLVHSHGILHARDLITSNLLITRNMIKTAADPLAQLQKMPHKHNTLCTPYIYP 180  
 QY 181 EIKTRSAHGLSDVMSLGCFYTLILGRPPDTIVKNTLNKVLAADYEMPSFLSTEARD 240  
 DB 181 EIKTRSAHGLSDVMSLGCFYTLILGRPPDTIVKNTLNKVLAADYEMPSFLSTEARD 240  
 QY 241 LIHQLLRRNPADRLSSSVLDHPFMSRNSYTSKDLGTEVSDISDGHATISYALTSAST 300  
 DB 241 LIHQLLRRNPADRLSSSVLDHPFMSRNSPKSKVGVTEVSDISDGHATISYALTSAST 300

QY 301 S1SGSLFDRRLIGOLPDKMTVF 325  
 DB 301 S1SGSLD-RLVLGOLPDKMTVF 324

RESULT 8  
 072VS3 PRELIMINARY; PRT; 711 AA.

AC 072VS3  
 DT 01-JUN-2003 (TEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Similar to serine/threonine kinase 18 (Fragment).  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Body;  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC045434; AAH45434.1; -  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007049; P:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; POLO box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR PROSITE; PS50078; POLO BOX; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR KINASE.  
 KW NON TER  
 FT NON TER  
 SQ SEQUENCE 711 AA; 77340 MW; B8326280CCE72EAC CRC64;

Query Match 27 %; Score 1408.5; DB 13; Length 711;  
 Best Local Similarity 43.3%; Pred. No. 1.3e-83;  
 Matches 334; Conservative 113; Mismatches 224; Indels 101; Gaps 24;

QY 230 MFSFLIEAKDLIQLRLRRPADRLSLVSLDHPHPSRNSSTKSLGTYE-DSISGHA 288  
 DB 1 MFMHISAEKODLIQLQLKPNPRLPSLVLDHPHTGOSPTAKSGSSGSSISGIA 60

QY 289 TISTATIASSTISGSLFDRRLIGOLPDKMTVF 345  
 DB 61 TISTASNATNNS-SSRLQRRTRQWIGOLPDRMAPIPSHSHSTSSSFKSGQD----- 113

QY 346 QW---GNGSTNSGRRGVQDAEE-RPHRYLRAVSSDRS--GTSNSGQAQTYTMC 399  
 DB 114 -QPNQSDLSKRGVGRIPVAGDSGRPHRYLRRASSDSVAVGSHNPQEA--LERC 169

QY 400 HSAEMLSVSKR-----SGGGENEERS-----PTDNANINFPFKETSSSSGSPER 446  
 DB 170 HSEEMLSGAGRLEPPQSGRYNAPHGYSKIDRLSPVVKQAPNASSF--STHSRQGM 227

QY 447 PNNQALSHLCPGKTPPPADPTPQETVQVQFGLQINALRKTEYDYSFN-RDPQ 505  
 DB 228 PQSQ-----TQWFSN-----DGVPKPADMS 249

QY 506 GHPDLDKQNSKNAWDTKYKNSDASDNHNSVQGMKMTALSKPELIOECYFGSD 565  
 DB 250 GHSSSGSFHSERGPITQT-SCSDKPSGLHS--QQQPIIFQ--HNPPGCRD-AFVSG 302

QY 566 PLSEQSKTRGME---PPM--GYQNR-----TLRSTSPVAHLRLPIROKTKKAVSII 614  
 DB 303 HNSEPPAYDAQPPCPPLSKGKXNTEKKXKVCCKSPPLCAARLPIROKTKKAVSII 362

QY 615 DSEEVCELVKEVYASGEYKEVLIQISSDQNTITTYVNGGGRFPLADRPSPPTDNISRY 674  
 DB 363 GNGEYCEMLIKGQAGQERVKEVLIQISSDQNTITTYVNGGGRFPLADRPSPPTDNISRY 422

QY 675 FDNLPKRYRKQYASRFPOLVRSKPKITTYFTRAKCILMENSPPADEFWFVGVKIH 734  
 DB 423 FDLPEKRYRKQYAKRFPOLVRSKPKITTYLTKPAKCHLMENSPNDELEVCDAKTH 482

QY 735 KTEP1QVIEKTKSYTLKSESEVNSLKKEIKMYMDHANEGRICLAL-STITSEERKT 793  
 DB 483 KTSQVRYVEKSGKSYTVAGDVGLSGINPECRILYIELSEGHMCLSLKALIAEGRSA 542

QY 794 RSPAPFPITIGRPGSTSSPKALSPPPVDSNPTPTDRASFRNMWNSAASPQAPILNP 853  
 DB 543 KNTFPFPITIGRPPVNPVPP---APAPSSSSC---RPAALAVHVLCSPPHOFITP 595

QY 854 SWNTNGLGLITTAAGTDISSLNKLCPKSAQLKSVFKVNGAATOLTSAGAWQVND 913  
 DB 596 SMISVAGSDLTTSVAKGSSPVNKHDERVTNSGVLKSLTPNIGWISQLTGSAWQVND 655

QY 914 GSQLVQVAGVSSISYTPNGQTRTYGENKLPDIYKQKLOCLSSIIILMFSNP 965  
 DB 656 GSQLVQVAGVSSISYTPNGQTRTYGENKLPDIYKQKLOCLSSIIILMFSNP 707

RESULT 9  
 097143 PRELIMINARY; PRT; 769 AA.

AC 097143  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE SAK protein.  
 GN SAK OR CG7186.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Phyllophaga; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.B., Gary N.S., Galbart W.M., Glaeser K.,  
 RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostali M., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,  
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclev J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hudson J.W., Dennis J.W.,  
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AEO03594; AAF51737.1; -  
 DR EMBL: AF106952; AAD19607.1; -  
 DR HSPSP: P00518; 1PHK  
 DR FlyBase: FBgn0026371; SAK.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004674; F:Protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0007049; P:cell cycle; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; P:kinase.  
 DR InterPro: IPR000719; P:kinase.  
 DR InterPro: IPR002290; Ser Thr kinase.  
 DR InterPro: IPR008266; Tyr kinase.  
 DR Pfam: PF00069; pkinase; 1  
 DR PRINTS: PR00109; TRYKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS0011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Kinase; transferase.  
 SQ SEQUENCE 769 AA; 85886 MW; F050BF60A5D94AA4 CRC64;

Query Match 26.4%; Score 1341.5; DB 5; length 769;  
 Best Local Similarity 33.0%; Pred. No. 3.6e-79;  
 Matches 321; Conservative 156; Mismatches 234; Indels 261; Gaps 24;

QY 6 GKEIDEPKGNLIGKSGFAGTVAESITGLEYALIMIDKKMYAGVQVQNEVKIHC 65  
 DB 8 GETTEDEYVQHLLKGGFATVVKARCIHTHODVAIMDKILQGTGLNVRQGEIHS 67  
 QY 66 QKHPSEILKLVNFFEDSNVYVLEWCHNGEMNRVYKKNVKEPSENEAHFHOITGM 125  
 DB 68 RLKHPVLOLTFPDQANVYVLELANHGEIARVYNNHARPTTEASILKQVAVGL 127  
 QY 126 YLHSHGILHDIYLSNLLITRNMTKIADFGLATQKMPHEKHYTLGTPNYISPEIATR 185  
 DB 128 YLHSHNIMHARDISLNLILSRMHYKIADFGLATQKMPDERHMTMCGTPNYSPEVVS 187  
 QY 186 SAHGESDVMSIGCMFYTLTIGRPEDDTVKTANKVYVLYADYEPSTLSIAKDLIHQL 245  
 DB 188 TSHGLPADWASVGMVYTLVGRPPETDAVOSTINKVMSRYIMPAHLSYEAQDLINKL 247  
 QY 246 LARNPADRLSSVLDHPFMSHNSSTKSKDLCTVEDSIDSGHATISTAITASSSTISGS 305  
 DB 248 LKGLPHERITLNAVLCHEPMLKCS-----NGHSA----- 277  
 QY 306 LFDKRLILGQPLVFNKATVFPNKKSTD-----FSSSGDGNSFYTQWQNETSNGRGV 360  
 DB 278 -----PALNVFSQSMESGDSGITTFASDSRSNQI-----RSEVNSGPGQV 320  
 QY 361 IODAEERPHSRYLRAVSSDRGTSNQSQAKTVMERCHAMLSVSKKGGGSENERK 420  
 DB 321 LPOIREERKQVHNKLPY----- 337  
 QY 421 SPTDNANITFNFPEKETSSTSSGSEFRPPNNOALSNHLCPKTPPPFADPTPQTEVQWF 480  
 DB 338 -----EQTGLF-----GQASGLAEP-----NWP 356  
 QY 481 GNLOINAHILKRTTEVDSISPRNDPGHDLQKOTSRYAATTTQYKQASDASDNAAHSVQO 540  
 DB 357 GAASSAPACMEA----- 368

QY 541 NTMKWTALHSKPEIIQOECVFGSDPLSEOSTKRGMEPPWQYQNRNLSITPLVAHRLK 600  
 DB 369 -----GNVPNSKQASL-----KEDHVSVPPLATYTKLL 395  
 QY 601 PIROKTKKAVNSILDSSEVCVAVK--EYASQERYKVLQISSDONTTITYPNGRGFP 658  
 DB 396 PTRYKTKNAIWSILNNGEVLAEFLKFRPTYNDRINDICRISDDCORIIITYOPDPGRGLP 455  
 QY 659 LADRPSS--PTDNIISRYSFNDLPEKYWRKYQYASRFQVLRKSKPKITYFRYAKTLM 715  
 DB 456 VREGPPDLQIPSGDCV--VNYDPLPSKMKKITYGARFGLVSKPKITYSTGLKCOLM 514  
 QY 716 EUSPGADREWVYDQVKTHT--EDFIQVIEKTKSYTLKSBEVNSLKEEIKMYDHANE 774  
 DB 515 EFM--TDEIFRYSAKILKTPSEGLKYVDRNGMLSDYSCESESRSL-----IEHNE 565  
 QY 775 GHRICTALESIISEBERKTRSAPEPILIGRPGSTSPKALSPPSVDSNYPRDRASF 834  
 DB 566 CETHCVNISMALEVAQTDNSC--PVTITGRPLTD-----VQPAQRIDG--LADTTN- 614  
 QY 835 NEMVMSAASPQAPILNPSMVTNGLGLTTASGTDISSNSLKDCLPKSAQ--LKSVE 892  
 DB 615 ---IAFSTPKNSQGSF--NFSLSL-----ISSTRMSTDEGTCNSMMAHQNIPIKRI 665  
 QY 893 VKNVGMATQVLSGAVWQVQNGSOLV---QAGVSSISTSPNGQTTREGENEKLDPYI 948  
 DB 666 VPEIDATLHSHGVQVQVQVGVSVVPSWQGG--GITYPNGTSTHFEGKGDLDLPFV 723  
 QY 949 KOKLQCLSSILL 960  
 DB 724 RDRVQQLINQL 735

RESULT 10  
 Q86NL8 PRELIMINARY; PRT; 769 AA.  
 AC Q86NL8;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE R570136p.  
 GN SAK.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Y;  
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Drenek D., Fafan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Munro J., Paclet J., Paragas V., Park S.,  
 RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,  
 RA Ceoliker S.;  
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: BT004646; AAC45202.1; -  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004674; F:Protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0007049; P:cell cycle; IEA.  
 DR InterPro: IPR000959; P:kinase.  
 DR InterPro: IPR000719; P:kinase.  
 DR InterPro: IPR002290; Ser Thr kinase.  
 DR InterPro: IPR001245; Tyr kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; Tyrc; 1.

DR PROSITE; PSS0078; POLO BOX; 1.  
 DR PROSITE; PSS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS00109; PROTEIN\_KINASE\_TYR; 1.  
 SQ SEQUENCE 769 AA; 85873 MW; E9241EFEB4FB90 CRC64;

Query Match 26.4%; Score 1338.5; DB 5; Length 769;  
 Best Local Similarity 32.9%; Pred. No. 5,7e-79;

Matches 320; Conservative 157; Mismatches 234; Indels 261; Gaps 24;

QY 6 GEKIEDPKVGNLLGKSGFAGVYRAESIHTGLKVAIKMIDKKMYKAGVQVQNEVKINC 65  
 DB 8 GETIEDVEYOHLLGKSGFATVYKARCLHTHODVAIKMIDKLLQGGLTRVYQVEIHS 67  
 QY 66 QLKHPSELILNYFEESNYYVLLVLEMCCHGEMNRILKAKYKPSSEAHKPHMQLITGM 125  
 DB 68 RLKHPPVLOLYTFQDANYYVLLVLELAHNGELHRYNNHILARPTELEASILQVAVGLL 127  
 QY 126 YLHSHGILHRDLTLNMLITRNANKIADFGLATOLKMPHEKHYTLCTPNYISPEIATR 185  
 DB 128 YLHSHIMHRDISLSNLYSREHNVKIDFGLATOLKRPERRIMTCPTNYISPEVSR 187  
 QY 186 SAHGESDVWSLQCMFYTLIGRPDPDTVNTKLVLYADYEMPSFLSEAKDLIHOL 245  
 DB 188 TSHGLPADVWSVQCMYTLIGRPDPDTVNTKLVLYADYEMPSFLSEAKDLIHOL 247  
 QY 246 LARNPADRLSLSVLDHPPMSRNSSTKSDLGTVESIDSGHATITAITASSSTISGS 305  
 DB 248 LKKLPHERITLVLVLCHPFLKCS-----NGGHA----- 277  
 QY 306 LFDKRLLLGQPLPNKWTVPKPKKSTSD-----FSSSGDINSFYTQMGNETSNGRGV 360  
 DB 278 -----FGLNLPFSQSMESGDSGITTFASDSNSQI-----RSVNSGPOV 320  
 QY 361 IODAEERPHSRYLKRAYSSDRGTSNSQOAKTYTWERCHSAEMLSVKRSGGGENERY 420  
 DB 321 LQDIREFQVHKLPLY----- 337  
 QY 421 SPFDNNANIPNFKEKTSSSSGSFERRDNNQALSNLCEKQTPPPADPPROTETVQWF 480  
 DB 338 -----BOTGLF-----GQASTGLAEP-----NMP 356  
 QY 481 GNLQIHAHLRKTTEYDISPNRDPQGHPLQKOTSXNMTDTLVKXNSADNAHSVKQ 540  
 DB 357 GAAKSAPFCMEA----- 368  
 QY 541 NTMKVYTAHLSKPEIIQOECVFGSDPLSEQSKTRGMEPPWGYQNRTLASTPLVAHLK 600  
 DB 369 -----GNVNSKQASL-----KEDRISVPLNTRKLL 395  
 QY 601 PIRQKTKAVVSTLDEEVCVELVYK--EYASQYVKEVLOISDQNTITTYNGRGPR 658  
 DB 396 PTRYKTKNIMSLIRNGEVVLEFLKFRPTNEDRINDICISIDGGRITIIYQDPGRGLP 455  
 QY 659 LADRPSS---PTDNIRYSFNDLPEKYMKRYOYASRFVOLVRSKRTYFTFYACILM 715  
 DB 456 VEEQPPDLQIPSDCV--YINDNLPKSKMKRTIYGAFVGLVSKTRPVYTFSLGCOLM 514  
 QY 716 ENSPGADFVWFYDVYKIHKT--EDPLOVIEKTKSYTLKSESEVNSLKEEIKMYMDHANE 774  
 DB 515 EFM--TDFEIRFYSGAKLKTPEBGLKVVDRNGMLSDVSCSRSRL-----IEKNGE 565  
 QY 775 GHRICLALSIISSEERKTRSAFPPIIIGRKRGSTSSPPLASPPSVDSNYPTRBRAS 834  
 DB 566 CFPTHCVNIALEVAQTKDNC--FPVITIGRRPITD-----VQPARLDD--LRDTTN-- 614  
 QY 835 NRVVMSAASPTQAPLINPBMVYNEGILTTTASGDTISNSLKDCLPKSAQ--LKSAYV 892  
 DB 615 ---IASSTPSKNGST--NFSLSL-----ISSTKPTSDPTGNCRSNMLAAHQIPIKRIN 665  
 QY 893 YKRVGATOLTSQAVVWQFNDGSQLVY---QAGVSIYSTSPNGQTRRYGENEKLPDYI 948  
 DB 666 VPEIGIATELSHGVVQVDFYDGSVSVIPMOG--GITVQPNGSTHSGKDDLPFPV 723

QY 949 KOKLOCLASSIL 960  
 DB 724 RDRVGOIPNITQL 735

RESULT 11

AC 080UT6 PRELIMINARY; PRT; 160 AA.  
 DT 01-JUN-2003 (TEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Similar to serine/threonine kinase 18 (Fragment).  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI TaxId=10090;

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Hellon E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.J., Skalek U., Smallos D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RA Strausberg R.,

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC051483; AAH51483.1; -

DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007049; P:cell cycle; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; POLO box.

DR PROSITE; PSS0078; POLO\_BOX; 1.  
 KW kinase

FT NON\_TER 1 1  
 SQ SEQUENCE 160 AA; 16579 MW; 8441CCIDA60B7238 CRC64;

Query Match 12.0%; Score 610; DB 11; Length 160;  
 Best Local Similarity 77.8%; Pred. No. 2,7e-32;  
 Matches 126; Conservative 15; Mismatches 15; Indels 6; Gaps 2;

QY 808 GSTSPKALSPPPVDSNYPTRDRAFNRMVMSAASPTQAPILNPSMNTNGLGITTTA 867  
 DB 4 GNTSSPKALSAAP--VDPSCCKGSDASRLSVNSAFAFPQSPGLSPSTVYVGLGHTATA 62  
 QY 868 SGTDISNSNLKQCLPSPADLKSVPKVGATOLTSQAVVWQFNDGSQLVQAGVSSIS 927  
 DB 63 TGTGVSSS-----LPSAQDLKSVFKNVGATOLTSQAVVWQFNDGSQLVQAGVSSIS 117  
 QY 928 YTSPPNGQTRRYGENEKLPDYIKOKLOCLASSILMPSNPTRNF 969

Db 118 YTPSDGQTRRYGENEKLEPIYIKOKLOCLSSIIIMFSNPTNPF 159

RESULT 12

086HN7 PRELIMINARY; PRT; 978 AA.

AC 086HN7; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to Aeterea pectinifera (Starfish). Polo-like kinase.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

NCBI\_Taxid=44689;

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RX MEDLINE=22092622; PubMed=12097910;

RA Gloeocher G., Bichinger U., Szafranski K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;

RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";

RL Nature 418:79-85(2002).

12]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA Baumgart C.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC116957; AAO52535.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0007049; P:cell cycle; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000959; POLO box.

DR InterPro; IPR000719; Prot kinase.

DR InterPro; IPR002290; Ser Thr\_pkinase.

DR InterPro; IPR001245; Tyr\_pkinase.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00210; S\_TKc; 1.

DR SMART; SM00219; TyrKc; 1.

DR PROSITE; PS00078; POLO BOX; 2.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS00113; PROTEIN KINASE\_DOM; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

KW Kinase.

SQ SEQUENCE 978 AA; 113599 MW; B03A677136543152 CRC64;

Query Match 11.7%; Score 593.5; DB 5; length 978;

Best Local Similarity 21.7%; Pred. No. 5e-30;

Matches 219; Conservative 166; Mismatches 303; Indels 323; Gaps 37;

QY 7 EKIEDFKVGNLKGKSFAGVRAESIHGELVAIKMIDPKAMVYKAGVQVONVEYKHCQ 66

DB 158 KILKEVROGEFLFGKGFACVLMTEVENRIRYAAKIIKSTLOKTRASKSEIKHS 217

QY 67 LKHPILBELVYFESNNVYVLENGCHGENRMYKKNVKPSENEAHFMHQTITGMLY 126

DB 218 LSHENIVKFEHCFENEENVYIILELNCQKTWMDIHKKR-KYLMVEYTKYVYUJIMAVQY 276

QY 127 LHSHGILHRDLTSLNLLTRNNNIIKADPGLATOLKMH-EKHYTLCTVPTVISELTR 185

DB 277 LHNNNIHRDLKGLNPLFD-NRITKADPGLST--KVEHGRKKITICTPYVIAPEIDN 333

QY 186 S-ANGLESDDVMSLGCMTYLLIGRPPFDITVKNLKNVLADEMP8--PLSIEADOLY 242

DB 334 SNGSYEVDVMSIGIILYLLIGKPPFSDVKHTYGRIKONQSFPEPIISHTGSLI 393

QY 243 HQLLRRNADLSSVLDHPFMSNSSTKSDQIGTVEDSIDSGHATITATASSTSI 302

DB 394 ISILNPVEORPNLTQILIEHDFYTSPIPK-----YLPVSSILTAPOSSTI 439

QY 303 SGLSFDKRLILGQPLPNMTVFPKNSSTDESSGDGNSFYTONGNETNSGGRVLYQ 362

DB 440 NQNM-----GRPLSEKTNIV-----NQQLQLAGTTSPTKNNHHYQ 476

QY 363 DAERPHSYRLRAVSSDSGTSNSQSAQKTYMERCHSAELSYSKSGGGENEERSP 422

DB 477 QYQQDPQOQY-----NNNYQSFSPKKQINNMMNNNNNNNNNNNNNNNNNN 523

QY 423 TNNNANFNPFKEKTSSSSGSFEPDPNQAALSNHLCPGKTPFPADPTPOTETVOQWFGN 482

DB 524 NNNNLLKQVY-----SNNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 564

QY 483 LQ-INAHLKTEYDISPNDPQGHPLQKDTSKNAWTDTRKYN-SDASDNHASYQKQ 540

DB 565 VENDDPHYRKLRLKEMEN-----DLK-----TOLLIKQYTNMNNNNNNNNNN 609

QY 541 NTMKTMTALHSPKEIIOQECVFGSDPLSEQSKTRGMEPPWGYQNRTLSSITPLVAHRLK 600

DB 610 -----QQQ----- 612

QY 601 PIRQTKKAAVSIIDSEEVCELVKEYASQBYKVELQISSDGNITITYPNGRGFP-- 658

DB 613 -----QQQQQQQQQFVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 648

QY 659 -----LADRPSPPTDNISRYSPDNLPKTKYKQYASRFVQLVRSKSPKITYPTR 708

DB 649 ELEETIANNNHISDSPEVSSNN-----NYP-----QQIQKQOP----- 680

QY 709 YAKCIIMENSPGADFEVWFYDG-----VKIKHTEDFIQLEKTKSYTLK----- 753

DB 681 -----NPNNEFYLGMPNNLVYISQVADF--TKYGLAVLVNSYGAVFND 724

QY 754 SESEVNSLKEIKMYMDHA-----NEGRICLALSIISBEERKTRSAAPF----- 799

DB 725 STKIYVLISEIAYMEHAKGTGDDGRVLANVTCQHPHDQKVTLLIKYLFHFTNSDTT 784

QY 800 PIIIGRKPGSTSPKALSPPSVDNYPTRDRASFNRMWMSAASPAQAILNPSWNTNE 859

DB 785 NLLI--NTGATSS-----SINN-----NNNNVEN- 807

QY 860 GLGLTTTASGTDISNSLKDCLPKSAQLKSVFK-----NVGMATQLTSGAAVWQFNDG 914

DB 808 -----VTNN 853

QY 915 SOLVWQGVSSISYTSNGQ-----TTRYGEN--EKLPDYIKQKLOCLSSIL 959

DB 854 SRIIV-SSKDMVTFVPRGQIITGTLNVPFGNGDKIKSEIKIYIGTILSNML 903

RESULT 13

Q7ZX26 PRELIMINARY; PRT; 663 AA.

AC 07ZX26; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to serum-inducible kinase.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

OC Xenopodidae; Xenopus.

NCBI\_Taxid=8355;

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Klein S., Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC045272; AA045272.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0007049; P:cell cycle; IEA.

DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002230; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR01245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO\_box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS50078; POLO\_BOX; 2.  
 DR PROSITE: PS50107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Kinase.  
 SQ SEQUENCE 663 AA; 76076 MW; 2E6BE9A37C37639C CRC64;  
 Query Match 11.7%; Score 592; DB 13; Length 663;  
 Best Local Similarity 33.1%; Pred. No. 3.5e-30;  
 Matches 119; Conservative 73; Mismatches 156; Indels 12; Gaps 4;  
 QY 15 GNLLGKGFAGYVRAESHTGLEVAIKMIDKKAMYGAVQVQVNEVKIHCOLKPSILE 74  
 DB 66 GKVLLGKGFARCYEMTDLTNNKIYAKIIPHSKVSXPHQREKIDKEIELHRTLHHRHVQ 125  
 QY 75 LNYFEDSNVYVLVLEMCHNGEMNRYLKNRYKPSNEARHFMHOITGMLYLHSHGILH 134  
 DB 126 FHHYFEDAKNIYIYLLEYCSRSSAHLTKR-KYLTPEVRYVYLKQIVSGIKTLHGEILH 184  
 QY 135 RDLTSLNLLTNNMIKIDFGLATOLKMPHEKHYTLGTPNYISPELATRAHGLESDV 194  
 DB 185 RDLKGNFPIITNMELKVGDFGLAARLEPEQRKRTICGTPNYLSBVLANKQHGESDI 244  
 QY 195 NSLGCMTYLLIGRPPTDVTYKNTLNKYLADYEMPSFLSIEAKDLIHQLRRPADRL 254  
 DB 245 MVLGCMYTMVLGSRPFETTLNLETYRCREARYSLPSSLMTSAKILASMLSRNEDRP 304  
 QY 255 SLSSVLDHFPMSNSSTKSLGTVEDSIDSGHATITATITASSSTISGSLPDKRLI 314  
 DB 305 HIDEIMQHDFTOGFTPERLPLCCHTADPHLSSPAKPFKKYAAALFGGKKEKSKYLD 364  
 QY 315 GQPLPKMTVPFKN-----KSTDFSSSGDNGSFYQWQNGQETNSGGRGVLDAREPH 369  
 DB 365 NH---NKL---PKDEEITKYLRODLQKTSISHQLNPRDDEIKNISKSDVLMKADKH 418  
 RESULT 14  
 Q96CV1 PRELIMINARY; PRT; 646 AA.  
 AC 096CV1; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DS Hypothetical protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]\_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: BC013899; AAH1389.1; -  
 DR GO:0005524; P:ATP binding; IEA.  
 DR GO:0006474; P:protein serine/threonine kinase activity; IEA.  
 DR GO:0016740; P:transferase activity; IEA.  
 DR GO:0007049; P:cell cycle; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR00719; Prot\_kinase.  
 DR InterPro: IPR002230; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR01245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO\_box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS50078; POLO\_BOX; 2.  
 DR PROSITE: PS50107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Kinase.  
 SQ SEQUENCE 663 AA; 76076 MW; 2E6BE9A37C37639C CRC64;

DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO\_box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS50078; POLO\_BOX; 2.  
 DR PROSITE: PS50107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Hypothetical protein; ATP-binding; Kinase;  
 KW serine/threonine-protein kinase; transferase.  
 SQ SEQUENCE 646 AA; 71628 MW; 3242ADD900865BE CRC64;  
 Query Match 11.6%; Score 590.5; DB 4; Length 646;  
 Best Local Similarity 37.5%; Pred. No. 4.2e-30;  
 Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;  
 QY 15 GNLLGKGFAGYVRAESHTGLEVAIKMIDKKAMYGAVQVQVNEVKIHCOLKPSILE 74  
 DB 65 GNLLGKGFARCYEMTDLTNNKIYAKIIPHSKVSXPHQREKIDKEIELHRTLHHRHVQ 124  
 QY 75 LNYFEDSNVYVLVLEMCHNGEMNRYLKNRYKPSNEARHFMHOITGMLYLHSHGILH 134  
 DB 125 FHHYFEDADNIYIYIELCSKSLAHYWKAR-HTLLEPEVRYVYLRLISGLKYLHOGILH 183  
 QY 135 RDLTSLNLLTNNMIKIDFGLATOLKMPHEKHYTLGTPNYISPELATRAHGLESDV 194  
 DB 184 RDLKGNFPIITNMELKVGDFGLAARLEPEQRKRTICGTPNYVAEVLRLGSGHEADV 243  
 QY 195 NSLGCMTYLLIGRPPTDVTYKNTLNKYLADYEMPSFLSIEAKDLIHQLRRPADRL 254  
 DB 244 MVLGCMYTMVLGSRPFETTLNLETYRCREARYSLPSSLMTSAKILASMLSRNEDRP 303  
 QY 255 SLSSVLDHFPMSNSSTKSLGTVEDSIDSGHATITATITASSST-----SISG 304  
 DB 304 SIDQLIHDFPK-----GYTPDRL-----DISCVTPDLPMPARSLFAKVK 349  
 QY 305 SLFDRK 311  
 DB 350 SLFGKX 356  
 RESULT 15  
 Q8K0J7 PRELIMINARY; PRT; 504 AA.  
 AC 08K0J7; 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DS Hypothetical protein.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RA Strausberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC031180; AAH31180.1; -  
 DR MGD; MG1:109604; Chk.  
 DR GO:0005524; P:ATP binding; IEA.  
 DR GO:0006474; P:protein serine/threonine kinase activity; IEA.  
 DR GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO:0016740; P:transferase activity; IEA.  
 DR GO:0007049; P:cell cycle; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR00719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.

DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00069; kinase; 1.  
DR Pfam; PF00659; POLO box; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS50076; POLO\_BOX; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KM Hypothetical protein: ATP-binding; Transferase.  
SQ SEQUENCE 504 AA; 55546 MW; FC4BD79BBD273DB9 CRC64;

Query Match 11.6%; Score 587.5; DB 11; Length 504;  
Best Local Similarity 36.8%; Pred. No. 4.5e-30;

Matches 113; Conservative 64; Mismatches 105; Indels 25; Gaps 4;

QY 15 GNLLGSGSPAGVYRAESIRHGLEVALKMDKKMYKAGVQVQNEVKIHQQLKPSILE 74  
DB 66 GRLLGKGGFARCYEATDTEGIAVAVKVIPOSRVAKPHQREKILNEIELHRDLQHRHIVR 125  
QY 75 LNYFEDSNVYVLEMCCHNGEMNRYLKVRKPFSENBAHFPHQIITGMLYHSHGILH 134  
DB 126 FSHHFFDADNIYFLELCGRKSLAHYKAR-HTLEPEVARYVDRQTLGKYLHQRGILH 184  
QY 135 RDIITLSNLLITRNMTIKIADFGIATQLKMPHEKHYTLCTPNYISPEIATRSNAGLESYV 194  
DB 185 RDLKLGNFPTIDNMEIKVGFGLARLEPPEQGRKTTICGTPNVVAPEVLLRQHGPEADV 244  
QY 195 WSIGCMFYTLIGRPPEDTDVTKNTLVKVVYADYEMPSFLSTEAKDLIHQLRRNPADRL 254  
DB 245 WSIGCMFYTLIGRPPEDTDVTKNTLVKVVYADYEMPSFLSTEAKDLIHQLRRNPADRL 304  
QY 255 SLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISTAITASSST-----SIG 304  
DB 305 SIEQILRHDFETK-----GYTPDRL-----PVSSCVTVPDLTTPNPASLPAKVTK 350  
QY 305 SLFEDKR 311  
DB 351 SLFGRKK 357

Search completed: September 23, 2004, 21:02:04  
Job time : 69 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 21:00:41 ; Search time 69 Seconds  
(without alignments)  
4520.463 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078  
Sequence: 1 MATCGEKIEDPKVGNLTKG.....KLQCLSSILMSPNTPNPH 970

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5078	100.0	970	14	US-10-026-021-2
2	5074	99.9	970	16	US-10-408-765A-1916
3	4861	95.7	928	12	US-10-425-114-37528
4	1973	38.9	379	14	US-10-026-021-3
5	590.5	11.6	607	10	US-09-769-970-15
6	590.5	11.6	607	14	US-10-108-580-2
7	590.5	11.6	607	14	US-10-204-041-16
8	590.5	11.6	607	16	US-10-620-052A-28
9	590.5	11.6	646	16	US-10-755-889-4
10	589.5	11.6	373	14	US-10-026-021-4
11	580.5	11.4	253	16	US-10-620-052A-43
12	568.5	11.2	1462	16	US-10-618-581-15
13	560.5	11.0	400	14	US-10-026-021-5
14	560.5	11.0	685	9	US-09-771-161A-249
15	560.5	11.0	685	9	US-09-771-161A-250

16	560.5	11.0	685	9	US-09-771-161A-251	Sequence 251, App
17	560.5	11.0	685	10	US-09-769-970-1	Sequence 1, Appl
18	560.5	11.0	685	12	US-10-260-708-69	Sequence 69, Appl
19	560.5	11.0	685	14	US-10-024-298A-101	Sequence 101, App
20	560.5	11.0	685	14	US-10-042-211A-101	Sequence 101, App
21	560.5	11.0	685	16	US-10-617-217A-101	Sequence 101, App
22	560.5	11.0	753	15	US-10-264-049-3124	Sequence 3124, App
23	544.5	10.7	1246	15	US-10-369-493-6585	Sequence 6585, App
24	538.5	10.6	896	16	US-10-437-963-136508	Sequence 136508, App
25	538	10.6	603	9	US-09-771-161A-214	Sequence 214, App
26	536	10.6	603	16	US-10-188-832-110	Sequence 110, App
27	536	10.6	603	16	US-10-408-765A-2279	Sequence 2279, App
28	536	10.6	629	12	US-10-425-114-37525	Sequence 37525, App
29	535.5	10.5	1038	12	US-10-424-599-274878	Sequence 274878, App
30	533	10.5	603	12	US-10-406-901-2	Sequence 2, Appl
31	533	10.5	603	14	US-10-171-311-186	Sequence 186, App
32	530	10.4	367	14	US-10-026-021-6	Sequence 6, Appl
33	527	10.4	461	16	US-10-767-701-36696	Sequence 36696, App
34	526.5	10.4	1518	9	US-09-801-368-152	Sequence 152, App
35	526.5	10.4	1518	15	US-10-369-493-22243	Sequence 22243, App
36	526.5	10.4	1518	16	US-10-618-581-11	Sequence 11, Appl
37	525	10.3	353	16	US-10-664-421-124	Sequence 124, App
38	524	10.3	523	12	US-10-183-687-258	Sequence 258, App
39	520	10.2	493	16	US-10-437-963-184443	Sequence 184443, App
40	517.5	10.2	1349	16	US-10-618-581-16	Sequence 16, Appl
41	516.5	10.2	329	9	US-09-925-300-1268	Sequence 1268, App
42	516.5	10.2	832	9	US-09-919-585-21	Sequence 21, Appl
43	516.5	10.2	832	15	US-10-464-939-2	Sequence 2, Appl
44	514.5	10.1	683	15	US-10-369-493-2188	Sequence 2188, App
45	510.5	10.1	402	12	US-10-087-684-51	Sequence 51, Appl

#### ALIGNMENTS

RESULT 1  
US-10-026-021-2  
; Sequence 2, Application US/10026021  
; Publication No. US20030027756A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoenhi, Yasumichi  
; APPLICANT: Demo, Susan  
; APPLICANT: Jenkins, Yonchu  
; APPLICANT: Rigel Pharmaceuticals, Inc.  
; TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
; FILE REFERENCE: 021044-001210US  
; CURRENT APPLICATION NUMBER: US/10/026,021  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/309,632  
; PRIOR FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 970  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human SAK serine/threonine kinase  
US-10-026-021-2

Query Match 100.0%; Score 5078; DB 14; Length 970;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 970; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATCGEKIEDPKVGNLTKGSPAGYRAESHTGTEVAIKNIIDKKAMTKAGVORVONE 60  
DB 1 MATCGEKIEDPKVGNLTKGSPAGYRAESHTGTEVAIKNIIDKKAMTKAGVORVONE 60  
QY VVHHCQLKRPSTLELYNYFEDSNYYVLVLEMGHGMARVYLNKRVKPFSENAAPHMHOI 120  
DB 61 VVHHCQLKRPSTLELYNYFEDSNYYVLVLEMGHGMARVYLNKRVKPFSENAAPHMHOI 120

```

QY 121 ITGMLYLHSHGILHRDLTLNSLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
Db 121 ITGMLYLHSHGILHRDLTLNSLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
QY 181 EIATRSAGHLSGSDVWSLGCMEFYTLILGRPPEDTDTVKNTLNKVLADYEMPSFLSTEAKD 240
Db 181 EIATRSAGHLSGSDVWSLGCMEFYTLILGRPPEDTDTVKNTLNKVLADYEMPSFLSTEAKD 240
QY 241 LHHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGVEDSLDSGHATITATITASSST 300
Db 241 LHHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGVEDSLDSGHATITATITASSST 300
QY 301 SISGSLFDRKRLILIGPLPNKMTVPFKXKSTDPSSSGDGNSEFYTOMGNOETSNISGRGV 360
Db 301 SISGSLFDRKRLILIGPLPNKMTVPFKXKSTDPSSSGDGNSEFYTOMGNOETSNISGRGV 360
QY 361 IODAEERPHSRYLRAVSSDRSGTNSOSQAKTYTMRCHSAEMLSVSKRSGGGENEERY 420
Db 361 IODAEERPHSRYLRAVSSDRSGTNSOSQAKTYTMRCHSAEMLSVSKRSGGGENEERY 420
QY 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
QY 481 GNLQINAHLRKTTEDYSISPNDPQGHPLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
Db 481 GNLQINAHLRKTTEDYSISPNDPQGHPLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
QY 541 NTKMKTALHSHKPEIIOQECVFGSDPLSFOSKTRGMEPWGQNRITLSITSPVAHRLK 600
Db 541 NTKMKTALHSHKPEIIOQECVFGSDPLSFOSKTRGMEPWGQNRITLSITSPVAHRLK 600
QY 601 PIROKTKKAVSILDSSEVCVELVKEVASQEVKVELQISSDGNITITYYPNKGRGFPILA 660
Db 601 PIROKTKKAVSILDSSEVCVELVKEVASQEVKVELQISSDGNITITYYPNKGRGFPILA 660
QY 661 DRPSPDNTISRYSFNDLPEKTYWKRYQYASRFVOLVRKSPKITTYFTRVAKCILMENSFG 720
Db 661 DRPSPDNTISRYSFNDLPEKTYWKRYQYASRFVOLVRKSPKITTYFTRVAKCILMENSFG 720
QY 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSSEVNSLKEEIKMWDHANEGRICL 780
Db 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSSEVNSLKEEIKMWDHANEGRICL 780
QY 781 ALESIISEBERKTSAPFPPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
Db 781 ALESIISEBERKTSAPFPPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
QY 841 SAASPTQAPILNPSVNTNGLGLTTTASGTDISSNSLKDCLEPKAQOLKSVFVKNVGNAT 900
Db 841 SAASPTQAPILNPSVNTNGLGLTTTASGTDISSNSLKDCLEPKAQOLKSVFVKNVGNAT 900
QY 901 QUTSAGVAVQFNDGSQLVVOAGVSSISYTSPNGOTTRRYGENEKLDPDYIKOKLQCLSSILL 960
Db 901 QUTSAGVAVQFNDGSQLVVOAGVSSISYTSPNGOTTRRYGENEKLDPDYIKOKLQCLSSILL 960
QY 961 MFSNPTPNFH 970
Db 961 MFSNPTPNFH 970

```

## RESULT 2

```

US-10-408-765A-1916
; Sequence 1916, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Rahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.

```

```

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1916
; LENGTH: 970
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1916

```

Query Match 99.9%; Score 5074; DB 16; Length 970;

Best Local Similarity 99.9%; Pred. No. 0; Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCIGEKIEDPKVGNLIGKSPAGVYRAESIHTGLJEVAIKMIDKKAMYKAGMYQRYONE 60
Db 1 MATCIGEKIEDPKVGNLIGKSPAGVYRAESIHTGLJEVAIKMIDKKAMYKAGMYQRYONE 60
QY 61 VKIHQLKHPSTILEYNYFEDSNYYVLYLWCHNGENRKYLNRYKPSSEBARHFHQI 120
Db 61 VKIHQLKHPSTILEYNYFEDSNYYVLYLWCHNGENRKYLNRYKPSSEBARHFHQI 120
QY 121 ITGMLYLHSHGILHRDLTLNSLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
Db 121 ITGMLYLHSHGILHRDLTLNSLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
QY 181 EIATRSAGHLSGSDVWSLGCMEFYTLILGRPPEDTDTVKNTLNKVLADYEMPSFLSTEAKD 240
Db 181 EIATRSAGHLSGSDVWSLGCMEFYTLILGRPPEDTDTVKNTLNKVLADYEMPSFLSTEAKD 240
QY 241 LHHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGVEDSLDSGHATITATITASSST 300
Db 241 LHHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGVEDSLDSGHATITATITASSST 300
QY 301 SISGSLFDRKRLILIGPLPNKMTVPFKXKSTDPSSSGDGNSEFYTOMGNOETSNISGRGV 360
Db 301 SISGSLFDRKRLILIGPLPNKMTVPFKXKSTDPSSSGDGNSEFYTOMGNOETSNISGRGV 360
QY 361 IODAEERPHSRYLRAVSSDRSGTNSOSQAKTYTMRCHSAEMLSVSKRSGGGENEERY 420
Db 361 IODAEERPHSRYLRAVSSDRSGTNSOSQAKTYTMRCHSAEMLSVSKRSGGGENEERY 420
QY 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
QY 481 GNLQINAHLRKTTEDYSISPNDPQGHPLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
Db 481 GNLQINAHLRKTTEDYSISPNDPQGHPLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
QY 541 NTKMKTALHSHKPEIIOQECVFGSDPLSFOSKTRGMEPWGQNRITLSITSPVAHRLK 600
Db 541 NTKMKTALHSHKPEIIOQECVFGSDPLSFOSKTRGMEPWGQNRITLSITSPVAHRLK 600
QY 601 PIROKTKKAVSILDSSEVCVELVKEVASQEVKVELQISSDGNITITYYPNKGRGFPILA 660
Db 601 PIROKTKKAVSILDSSEVCVELVKEVASQEVKVELQISSDGNITITYYPNKGRGFPILA 660
QY 661 DRPSPDNTISRYSFNDLPEKTYWKRYQYASRFVOLVRKSPKITTYFTRVAKCILMENSFG 720
Db 661 DRPSPDNTISRYSFNDLPEKTYWKRYQYASRFVOLVRKSPKITTYFTRVAKCILMENSFG 720
QY 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSSEVNSLKEEIKMWDHANEGRICL 780
Db 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSSEVNSLKEEIKMWDHANEGRICL 780
QY 781 ALESIISEBERKTSAPFPPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
Db 781 ALESIISEBERKTSAPFPPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840

```

QY 841 SAASPTQADILNPNMVTNEGILGTTTASGTDISSNLSKDCLPKSAQLLKSFVKNVGMAT 900  
 DB 841 SAASPTQADILNPNMVTNEGILGTTTASGTDISSNLSKDCLPKSAQLLKSFVKNVGMAT 900  
 QY 901 QLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENEXLPDYIKQKQCLSSILL 960  
 DB 901 QLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENEXLPDYIKQKQCLSSILL 960  
 QY 961 MFSNPTNPFH 970  
 DB 961 MFSNPTNPFH 970

### RESULT 3

US-10-425-114-37528  
 / Sequence 37528, Application US/10425114  
 / Publication No. US20040034888A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Lili, Jindong  
 / APPLICANT: Zhou, Yihua  
 / APPLICANT: Kovalic, David K.  
 / APPLICANT: Screen, Steven E.  
 / APPLICANT: Tabaska, Jack E.  
 / APPLICANT: Cao, Yongwei  
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 / FILE REFERENCE: 38-21(5313) B  
 / CURRENT APPLICATION NUMBER: US/10/425, 114  
 / CURRENT FILING DATE: 2003-04-28  
 / NUMBER OF SEQ ID NOS: 73128  
 / SEQ ID NO 37528  
 / LENGTH: 928  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / OTHER INFORMATION: Clone ID: LIB4119-112-84\_FRI.pep  
 / US-10-425-114-37528

Query Match 95.7%; Score 4861; DB 12; Length 928;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 927; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 43 IDKKAMVKAQMVQVONEVYKHCQLKHPSTILELYNFBDSNYVYLLEMCHNEMKRYLK 102  
 DB 1 IDKKAMVKAQMVQVONEVYKHCQLKHPSTILELYNFBDSNYVYLLEMCHNEMKRYLK 60  
 QY 103 NRVKPESEBARHFMHQITGMVYLHSHGILHRDLTSLNLLTRNNVIXIADFGLATOLK 162  
 DB 61 NRVKPESEBARHFMHQITGMVYLHSHGILHRDLTSLNLLTRNNVIXIADFGLATOLK 120  
 QY 163 MPHEKAYTLCGPPNYISPELATRSANGLESVDWVSGCMFTYLLIGRPPPTDTVKNYTLNK 222  
 DB 121 MPHEKAYTLCGPPNYISPELATRSANGLESVDWVSGCMFTYLLIGRPPPTDTVKNYTLNK 180  
 QY 223 VVLADYEMPSFISIEKDLIHLRRNPADRLSLSSVLDHPFMRSSSTSKDGLTVEDS 282  
 DB 181 VVLADYEMPSFISIEKDLIHLRRNPADRLSLSSVLDHPFMRSSSTSKDGLTVEDS 240  
 QY 283 IDSGHATITATITASSSTSISGLPDKRRLIGQPLPNKATVPPKXKSTDFSSSGDGNS 342  
 DB 241 IDSGHATITATITASSSTSISGLPDKRRLIGQPLPNKATVPPKXKSTDFSSSGDGNS 300  
 QY 343 FYTQGNQSTNSGGRGVIODAERPHSRYLRAVYSSDNGSGTNSGSOAKTYTMEKCHA 402  
 DB 301 FYTQGNQSTNSGGRGVIODAERPHSRYLRAVYSSDNGSGTNSGSOAKTYTMEKCHA 360  
 QY 403 EMLSVKSGGGGNEERYSPDTNNANIFNFKEKTSSSSGSPRPNNQOLSHLCPGKT 462  
 DB 361 EMLSVKSGGGGNEERYSPDTNNANIFNFKEKTSSSSGSPRPNNQOLSHLCPGKT 420  
 QY 463 PPPFADPTPTETVVOQFGLQINAHLRKTEYDSSISPNRDFQGHDLQKDTSKNMTDT 522  
 DB 421 PPPFADPTPTETVVOQFGLQINAHLRKTEYDSSISPNRDFQGHDLQKDTSKNMTDT 480

QY 523 KYKNSDASDNASHVYKQNTMKMTALHSEKPEIIQOECVFGSDPLSEQSKTRGMEPPMGY 582  
 DB 481 KYKNSDASDNASHVYKQNTMKMTALHSEKPEIIQOECVFGSDPLSEQSKTRGMEPPMGY 540  
 QY 583 QNRTLSITSPLVAHRLKPIROKTKKAVVSIIDSEEVCELVKEYASQEVYXEVLIQISSD 642  
 DB 541 QNRTLSITSPLVAHRLKPIROKTKKAVVSIIDSEEVCELVKEYASQEVYXEVLIQISSD 600  
 QY 643 GNTTITTYPNGRGFPPLADRPSPPTDNIISRYSPDNIPEKYMRYQYASFPVLVNSKSK 702  
 DB 601 GNTTITTYPNGRGFPPLADRPSPPTDNIISRYSPDNIPEKYMRYQYASFPVLVNSKSK 660  
 QY 703 ITYPTRYACIIMENSPGADPEWAFYDGYKIKHTEDPIOVIRKTSKSYLKSESEVNSLK 762  
 DB 661 ITYPTRYACIIMENSPGADPEWAFYDGYKIKHTEDPIOVIRKTSKSYLKSESEVNSLK 720  
 QY 763 BEIKMYMDHANGHRICLALESIISEEKRTSAPFPPIIGRKPGSTSPKALSPPSV 822  
 DB 721 BEIKMYMDHANGHRICLALESIISEEKRTSAPFPPIIGRKPGSTSPKALSPPSV 780  
 QY 823 DSNYPTRDASFRKMYMASHAAPTQAPILNPMVTNEGILGTTTASGTDISSNLSKDCLP 882  
 DB 781 DSNYPTRDASFRKMYMASHAAPTQAPILNPMVTNEGILGTTTASGTDISSNLSKDCLP 840  
 QY 883 KSAQLLKSFVKNVGMATQLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENE 942  
 DB 841 KSAQLLKSFVKNVGMATQLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENE 900  
 QY 943 KLPDYIKQKQCLSSILLMFSNPTNPFH 970  
 DB 901 KLPDYIKQKQCLSSILLMFSNPTNPFH 928

### RESULT 4

US-10-026-021-3  
 / Sequence 3, Application US/10026021  
 / Publication No. US20030027756A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Hiteoshi, Yasumichi  
 / APPLICANT: Demo, Susan  
 / APPLICANT: Jenkins, Yonchu  
 / APPLICANT: Rigel Pharmaceuticals, Inc.  
 / TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
 / TITLE OF INVENTION: Treatment of Cancer  
 / FILE REFERENCE: 021044-001210US  
 / CURRENT APPLICATION NUMBER: US/10/026, 021  
 / CURRENT FILING DATE: 2002-06-25  
 / PRIOR APPLICATION NUMBER: US 60/309, 632  
 / PRIOR FILING DATE: 2001-08-01  
 / NUMBER OF SEQ ID NOS: 8  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 3  
 / LENGTH: 379  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: DOMAIN  
 / LOCATION: (1)..(379)  
 / OTHER INFORMATION: SAK serine/threonine kinase domain  
 / US-10-026-021-3

Query Match 38.9%; Score 1973; DB 14; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-127;  
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATCIEKIEDPFGVNLGKGSFAGYVRAESIHTGLEVAIKIIDKKAMVKAQVQVONE 60  
 DB 1 MATCIEKIEDPFGVNLGKGSFAGYVRAESIHTGLEVAIKIIDKKAMVKAQVQVONE 60  
 QY 61 VKIHQCLKHPSILELYNFBDSNYVYLLEMCHNEMKRYLKRVKPESEBARHFMHQI 120  
 DB 61 VKIHQCLKHPSILELYNFBDSNYVYLLEMCHNEMKRYLKRVKPESEBARHFMHQI 120

Qy	121	ITGMVLKHSIGLHRLTLTNNLIRNNMIKADFGALTOLKMPHEKHYTLCTGPNTISP	180
Db	121	ITGMVLKHSIGLHRLTLTNNLIRNNMIKADFGALTOLKMPHEKHYTLCTGPNTISP	180
Qy	121	ITGMVLKHSIGLHRLTLTNNLIRNNMIKADFGALTOLKMPHEKHYTLCTGPNTISP	180
Db	121	ITGMVLKHSIGLHRLTLTNNLIRNNMIKADFGALTOLKMPHEKHYTLCTGPNTISP	180
Qy	181	EITRSAGHGESDWSLGCMPFTLLIGRPEDTDVTKNTLKNVLADEMPSPISIAKD	240
Db	181	EITRSAGHGESDWSLGCMPFTLLIGRPEDTDVTKNTLKNVLADEMPSPISIAKD	240
Qy	241	LIRQLLRPNPADLSLSVLDHPFMRSNSTSKSLGTVEDSIDSGHATLSTATTAASST	300
Db	241	LIRQLLRPNPADLSLSVLDHPFMRSNSTSKSLGTVEDSIDSGHATLSTATTAASST	300
Qy	301	SISGSLFDKRRLLIGPPLFNKOTVPPKXKSTDFSSGDGNSFYTOWGNDETNSNGRIV	360
Db	301	SISGSLFDKRRLLIGPPLFNKOTVPPKXKSTDFSSGDGNSFYTOWGNDETNSNGRIV	360
Qy	361	IODAERPHSRYYLRAYSS	379
Db	361	IODAERPHSRYYLRAYSS	379

```

1      RESULT 5
2      US-09-769-970-15
3      ? Sequence 15, Application US/09769970
4      ? Publication No. US20030170219A1
5      ?
6      ? GENERAL INFORMATION:
7      ?
8      ? APPLICANT: Bandman, Olga
9      ? Hillman, Jennifer L.
10     ? Corley, Neil C.
11     ? Guegler, Karl G.
12     ? Lai, Preeti
13     ? Goli, Surya K.
14     ? Shah, Purvi
15     ?
16     ? TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
17     ? KINASES
18     ?
19     ? NUMBER OF SEQUENCES: 21
20     ?
21     ? CORRESPONDENCE ADDRESSES:
22     ? ADDRESS: Incyte Pharmaceuticals, Inc.
23     ? STEEHL: 3174 Porter Drive
24     ? CITY: Palo Alto
25     ? STATE: CA
26     ? COUNTRY: USA
27     ? ZIP: 94304
28     ?
29     ? COMPUTER READABLE FORM:
30     ? MEDIUM TYPE: Diskette
31     ? COMPUTER: IBM Compatible
32     ? OPERATING SYSTEM: DOS
33     ? SOFTWARE: FastSeq for Windows Version 2.0.
34     ?
35     ? CURRENT APPLICATION DATA:
36     ? APPLICATION NUMBER: US/09/769,970
37     ? FILING DATE: 24-Jan-2001
38     ? CLASSIFICATION: <Unknown>
39     ?
40     ? PRIOR APPLICATION DATA:
41     ? APPLICATION NUMBER: 09/272,796
42     ? FILING DATE: <Unknown>
43     ?
44     ? ATTORNEY/AGENT INFORMATION:
45     ? NAME: Billings, Lucy J J
46     ? REGISTRATION NUMBER: 36,749
47     ? REFERENCE/DOCKET NUMBER: PF-0321 US
48     ? TELECOMMUNICATION INFORMATION:
49     ? TELEPHONE: 415-855-0555
50     ? TELEFAX: 415-845-4166
51     ? TELEX: <Unknown>
52     ?
53     ? INFORMATION FOR SEQ ID NO: 15:
54     ? SEQUENCE CHARACTERISTICS:
55     ? LENGTH: 607 amino acids
56     ? TYPE: amino acid
57     ? STRANDEDNESS: single
58     ? TOPOLOGY: linear
59     ?
60     ? IMMEDIATE SOURCE:
61     ?
62     ? LIBRARY: GenBank
63     ? CLONE: 1827450

```

SEQUENCE DESCRIPTION:	SEQ ID NO:	15:
US-09-769-970-15		

	Query Match	11.6%;	Score 590.5;	DB 10;	Length 607;	
	Best Local Similarity	37.5%;	Pred. No. 7,1e-32;			
	Matches	115;	Conservative	61;	Mismatches 106;	Indels 25; Gaps 4
QY	15	GNIIGKGSFAGVYRAESIHITGLEVAIAIKMIDKKAYKACMVORVONEVIHICQLKPSILE	74			
Db	26	GRLLGKGGFACCYEATDTETSSAIVAVKIIPQSRAVKPHQREKLINIEIHRDLQRHHIVR	85			
QY	75	LNYTFEDSNYYVLVLEMCHEBNMRYLKNRYKPSENAARHPMDIITGMILYHSHGITLH	134			
Db	86	FSHFPEADNITYIFLELCSRKSIAHIWVAR-HTLLEPRVARYLLRGLSGLKTIHQRGILH	144			
QY	135	RDLTFLSNILLTRNNNIKIDAGLGATOLMKPHEKYTLTGTPNYISPELATRSAHGLSPDY	194			
Db	145	RDCLKGNPFITENWEIKVDGRLARLRPEPQRKKTICTGTNYAAVEYLKQHGHPEADV	204			
QY	195	WSLGCMEYTLIIGRPFPDPTDVTKNTLNKVLYADYEMPSFLISEAKDLTHOILRRNPADR.L	254			
Db	205	WSLGCVMYTLIICTGGPPFETADLKETGYRCIKGVHTPLPASLSIPAQGLLAAILRASPRDRP	264			
QY	255	SLSGVLDHPMSNRNSSTSKDLGTVEDSIDSGHAITSALTAASST-----SISG 304				
Db	265	SIIDLIHLHDFFTK-----GYTPRRL-----PISSCIVVPDULTPPNPARSLFAKYTK 310				
QY	305	SLPDKRR 311				
Db	311	SLFGKK 317				

```

RESULT 6
US-10-108-580-2
; Sequence 2, Application US/10108580
; Publication NO. US20030077681A1
; GENERAL INFORMATION:
; APPLICANT: Cogswell, John
; TITLE OF INVENTION: PKI3 PROTEIN-PROTEIN INTERACTIONS
; FILE REFERENCE: P04458
; CURRENT APPLICATION NUMBER: US/10/108,580
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 607
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-108-580-2

```

	Query Match	Best Local Similarity	11.6%;	Score 590..5;	DB 14;	Length 607;
	Matches	115;	Conservative	61;	Mismatches	106;
					Indels	25;
					Gaps	4;
QY	15	GNLLGKSPAGVYRAESIHTGLVAITKIDKKMYKAGMORVQNEVYKIHQOLKRPISLE	74			
DB	26	GRLLGKGGFARCYEATDTETGSAVAAYKVIPOSRAPKPHQRKILNELRLHDLQDRHIVR	85			
QY	75	LYNYFEDSNVYVLTLEMCHNGENMYUKNRPKPSSENEARHFHQIITGMILYHSHGILH	134			
DB	86	FSHHFEDADNIVYIFLELCSRKSLAHMKAR..HTLLEPVRVYLLQIISGLTKYLIHQRGILH	144			
QY	135	RDLTLSMLLTIRNNITKIADPGLATOLKMPHEKITYTLCGTPNTYISPELATSAGLESDV	194			
DB	145	RDILKGNFTTENNELKVGDFGLAARLEPREQRKTTICGPNVYAPEVLLRQGGPREADV	204			
QY	195	WSLGCMEFYTLIGRPEDTDTVKNLTINKVVLADAYEMSFSLIAKDLIHQLRBNPADRL	254			
DB	205	WSLGCVMYTLTLCSPPEFADIKETRYACIQVHNTLRASLSPARQLLAALTRASPRRR	264			
QY	255	SLSSVLDHPMSRNRSTKSKDLGTVEQSDISGNAITITATYASSST-----SISG	304			
DB	265	SIQQLIHDFFTK-----GTPPRL-----PISSCVVPDLTPNPASLFAVYTK	310			

QY 305 SLFDRR 311  
DB 311 SLFGRK 317

# RESULT 7

US-10-204-041-16  
Sequence 16, Application US/10204041  
Publication No. US20030176443A1  
GENERAL INFORMATION:  
APPLICANT: STEIN-GERLACH, MATTHIAS  
APPLICANT: SALASIDIS, KONSTANTINOS  
APPLICANT: BACHER, GERALD  
APPLICANT: MULLER, STERAN  
TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Prio  
FILE REFERENCE: AXM-007.1P US  
CURRENT APPLICATION NUMBER: US/10/204,041  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: EP 0111858.5  
PRIOR FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: PCT/EP02/05420  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 607  
TYPE: PR  
ORGANISM: Homo sapiens  
US-10-204-041-16

Query Match 11.6%; Score 590.5; DB 14; Length 607;  
Best Local Similarity 37.5%; Pred. No. 7.1e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 74  
DB 26 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 85  
QY 75 LNYFEDSNVYLVLEMCHNGEMNRYLKORVKPFSENEARHFMHQITITGMLYLSHGILH 134  
DB 86 FSHHFEADNIIYFELGSKSLAHIMWKAR-HTLLEPVRVYLRQLLSGLKYLHGQGIH 144  
QY 135 RDLTSLNLLTTRNNIKIADFGIATOLKMPHEKHYTLCTGPNYISPEIATRAHGLESDV 194  
DB 145 RDLKGNFETITMEIKVDFGLAARLEBPBQRKKTICGPNYVABEVLLRQGHGEADV 204  
QY 195 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 254  
DB 205 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 264  
QY 255 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
DB 265 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
QY 305 SLFDRR 311  
DB 311 SLFGRK 317

# RESULT 8

US-10-620-052A-28  
Sequence 28, Application US/10620052A  
Publication No. US20040126784A1  
GENERAL INFORMATION:  
APPLICANT: Hitoebl, Yasumichi  
APPLICANT: Jenkine, Yonchu  
APPLICANT: Markovtsov, Vadim  
APPLICANT: Rigel Pharmaceuticals, Inc.  
TITLE OF INVENTION: Modulators of Cellular Proliferation  
FILE REFERENCE: 021044-004010US  
CURRENT APPLICATION NUMBER: US/10/620,052A

CURRENT FILING DATE: 2003-07-14  
PRIOR APPLICATION NUMBER: US 60/395,443  
PRIOR FILING DATE: 2002-07-12  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 607  
TYPE: PR  
ORGANISM: Homo sapiens

OTHER INFORMATION: cytokine-inducible kinase (CKK) serine threonine  
OTHER INFORMATION: kinase, proliferation-related kinase (PRK),  
OTHER INFORMATION: polo-like kinase 3 (PLK3)  
US-10-620-052A-28

Query Match 11.6%; Score 590.5; DB 16; Length 607;  
Best Local Similarity 37.5%; Pred. No. 7.1e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 74  
DB 26 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 85  
QY 75 LNYFEDSNVYLVLEMCHNGEMNRYLKORVKPFSENEARHFMHQITITGMLYLSHGILH 134  
DB 86 FSHHFEADNIIYFELGSKSLAHIMWKAR-HTLLEPVRVYLRQLLSGLKYLHGQGIH 144  
QY 135 RDLTSLNLLTTRNNIKIADFGIATOLKMPHEKHYTLCTGPNYISPEIATRAHGLESDV 194  
DB 145 RDLKGNFETITMEIKVDFGLAARLEBPBQRKKTICGPNYVABEVLLRQGHGEADV 204  
QY 195 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 254  
DB 205 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 264  
QY 255 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
DB 265 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
QY 305 SLFDRR 311  
DB 311 SLFGRK 317

# RESULT 9

US-10-755-889-4  
Sequence 4, Application US/10755889  
Publication No. US20040171823A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
FILE REFERENCE: P4784 NP  
CURRENT APPLICATION NUMBER: US/10/755,889  
CURRENT FILING DATE: 2004-01-13  
PRIOR APPLICATION NUMBER: U.S. 60/440,068  
PRIOR FILING DATE: 2003-01-14  
PRIOR APPLICATION NUMBER: U.S. 60/469,757  
NUMBER OF SEQ ID NOS: 823  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 4  
LENGTH: 646  
TYPE: PR  
ORGANISM: Homo sapiens  
US-10-755-889-4

Query Match 11.6%; Score 590.5; DB 16; Length 646;  
Best Local Similarity 37.5%; Pred. No. 7.7e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 74  
DB 26 GNLLGKSPAGVYRAESHTGLEVAIKMTDKKAMYAGWVORVONEVKIHCOLKHPISLE 85  
QY 75 LNYFEDSNVYLVLEMCHNGEMNRYLKORVKPFSENEARHFMHQITITGMLYLSHGILH 134  
DB 86 FSHHFEADNIIYFELGSKSLAHIMWKAR-HTLLEPVRVYLRQLLSGLKYLHGQGIH 144  
QY 135 RDLTSLNLLTTRNNIKIADFGIATOLKMPHEKHYTLCTGPNYISPEIATRAHGLESDV 194  
DB 145 RDLKGNFETITMEIKVDFGLAARLEBPBQRKKTICGPNYVABEVLLRQGHGEADV 204  
QY 195 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 254  
DB 205 NSLGCWFYTLIGRPFDVDTVNTLKNVVLADYEMPSFLISAKDLIHQLLRNPADRL 264  
QY 255 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
DB 265 SLSSVLDHPFMSRNSSTKSKDLCTVEDSIDSGHATISTATSSST-----SISG 304  
QY 305 SLFDRR 311  
DB 311 SLFGRK 317

```

Db      65  GRLLGGSGPARCEAEDTDTETGSAVAVKVYIPOSRAVAKPHQREKILNIEIHRDILQHRIIVR 124
QY      75  LKYYEDSNYYVLYVLEMGCHGEMNRVYLRKRVKPSBENARHFGQIITGMLYLHSHGILH 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      125  FSHHFEEDANNTYIFELCSRKSLAHIMWR -HTLLEPVRVRYLLQIISGLKYLHQRGILH 183
QY      135  RDVLTLSNLLITRNMKIKIDFGIATOLMKPMHEKHVTLGTEPNYISPEIATRSAHGLESDV 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      184  RDLKLGNFPTTEMELKVGDFGIARLRLEPBGQRKKTICGTNNYVAPVYLRLQGHGREADV 243
QY      195  WSLGCMFPTLLIGRPPEFDTVKNTLANKVLYADYEMSPFLSTEAKDLIHQLRNPDRDL 254
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      244  WSLGCMVYVTLGSPPEFTADIKETRYCKYKOUHYVLPASLISLPARQLAALIRASPERDR 303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      255  SLSSVLDHFEFMSNBSYTSKDLGTVEDSIDSGHATITATYASSST-----SISG 304
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      304  SIDQILRHDFPK-----GYPRDL-----PLSSCYVVPDLTPPNPARSLFAKVTK 349
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      305  SLPEDKR 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      350  SLFGRRK 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
US-10-026-021-4
; Sequence 4, Application US/10026021
; Publication No. US20030027756A1
; GENERAL INFORMATION:
; APPLICANT: Hitoshi, Yasumichi
; APPLICANT: Demo, Susan
; APPLICANT: Jenkins, Yonchu
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: SA: Modulation of Cellular Proliferation for
; TITLE OF INVENTION: Treatment of Cancer
; FILE REFERENCE: 021044-001210US
; CURRENT APPLICATION NUMBER: US/10/026,021
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/309,632
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(373)
; OTHER INFORMATION: human FNK mitotic kinase domain
US-10-026-021-4

```

Query Match	11.6%	Score 589.5	DB 14	Length 373
Similarity	37.5%	Pred. No. 4:28-32		
Best Local		61	Mismatches 106	Indels 25
Matches 115				Gaps 4
Conservative				

```

QY      1 GNLDGSGSAGYARRESHTTGLEVAIKMIDKKAMVKKAMORVONEKHCIOLEKHPSEILE  74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      65 GMLDGGGCFARCYEATMDTETSAAVVKYIPSRVYVKHQNEKILNLEIEMRDLQHNRYA  124
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      75 LNNYPEDSNVYLVLEMCNGEMNRVYLKORVAKPSENEARHFMFOITGMYLIYHSHGILL  134
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      125 FSHHFDADANITYIELECSRSLAHIWVAR-HTLLEPPVRVRYQLILSGAKTYHQRILL  183
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      135 RDLTSLNLLTNNNNIKIADFGLATOLKMPHEKHVYLGCPYNTISPEIARSHAGIESDV  194
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      184 RDLKGNPFITENMELKVGDFGLARLRAPPORCKKTIGTFPNVYAPFVILROGHGEADV  243
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      195 WSLGCMFYLLIGRPPDPDYKNTLNKVLADVEMPSFISIEKDLIHOILARRNPADRL  254
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      244 WSLGCMYMTLLCGSPPEFADIKEYICIKQVHTTPASLSLPRKQMLAAILRASPDRP  303
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      255 SLSSVVLDPFMSRNSSTYSKDLGTVEDSISGHAITSTAITASSST-----SIGG  304
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

Db      304 SLDDIIRHDFYTK-----GYTPDKL-----PISCVTVPLTPPNPARSLFAKYTK 349
QY      305 SLFDKRR 311
        ||| :::
Db      350 SLFVRKK 356

```

RESULT 11  
 US-10-620-052A-43  
 Sequence 43, Application US/10620052A  
 Publication No. US20040126784A1  
 GENERAL INFORMATION:  
 APPLICANT: Hitoshi, Yasumichi  
 APPLICANT: Jenkins, Yonchu  
 APPLICANT: Markovtsov, Vadim  
 APPLICANT: Rigel Pharmaceuticals, Inc.  
 TITLE OF INVENTION: Modulators of Cellular Proliferation  
 FILE REFERENCE: 021044-004010US  
 CURRENT APPLICATION NUMBER: US/10/620, 052A  
 CURRENT FILING DATE: 2003-07-14  
 PRIOR APPLICATION NUMBER: US 60/395,443  
 PRIOR FILING DATE: 2002-07-12  
 NUMBER OF SEQ ID NOS: 78  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 43  
 LENGTH: 253  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: cytokine-inducible kinase (CNK) serine threonine  
 OTHER INFORMATION: kinase, proliferation-related kinase (PRK),  
 OTHER INFORMATION: polo-like kinase 3 (PLK3)  
 US-10-620-052A-43

Query Match	11.4%	Score 580.5	DB 16	Length 253
Best Local Similarity	42.4%	Pred. No. 1e-31		
Matches 106, Conservative	51	Mismatches	92	Gaps 1

[illegible]

```

RESULT 12
US-10-618-581-15
; Sequence 15, Application US/10618581
; Publication No. US20040077524A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, GERRAD
; APPLICANT: SUBDRASANA, SUCHA
; TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITING FUNGAL GROWTH
; FILE REFERENCE: 034536/0923
; CURRENT APPLICATION NUMBER: US/10/618,581
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,624
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 42

```

SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO 15  
 LENGTH: 1462  
 TYPE: PRT  
 ORGANISM: Candida albicans  
 US-10-618-581-15

Query Match 11.2%; Score 568.5; DB 16; Length 1462;  
 Best Local Similarity 26.5%; Pred. No. 7.8e-30;  
 Matches 189; Conservative 107; Mismatches 255; Indels 163; Gaps 21;

QY 8 KIEDPKVGNLKGSGFAGYRAESIHITGLEVAIKMIDKAMK-----AGNV 54  
 DB 61 KIGPKLGRGTGRSGRVLAKNTTGGDLAVKIYKSNPKLEPKYKRSKEDATRLP 120  
 QY 55 QRVQNEVKIKHCOLKPSILELYNYPEDSNVYVLVLEMCHNGEMNRYLKNRVKPFSENEK 114  
 DB 121 YGIREELIIMKLIISHNIMGLYDVWENKNDLYLLEYIGGELFDYLIRGK-LOEYEA 179  
 QY 115 HPMHQLITGMLYHSHGILHRDLTSLNLLTTRNMTKIDFGLATOLKMPHEKHYTLGT 174  
 DB 180 NFKCQIINGINYLHGFNICHRLKPEMLLDFKCNKIADFGMAA-LEVKEKLETSQGS 238  
 QY 175 PNYISPEI-ATRSAGLESVDVSLGCMFYTLIGRPFPDPTVKNLANKVVLADYEMPSF 233  
 DB 239 PNYASPEIYAGKNYHGAPESDIWSGILFLALLTGHLPFDENIRKLLKLVQSGKFMPE 298  
 QY 234 LIEAKDLIHLRLRNPAURLSLSVLDHPFMSR-----NSSTYSKDLGYE---DS 282  
 DB 299 LSFPEAKDLITKMLKVPREPRITIDALITPLAKYDEPTVSYSTTLTLDINSINIKOIS 358  
 QY 283 IDSGLATITAT-----ASSSTISGSLPDKRILLIGPLNPKMTVPFKNK-----S 330  
 DB 359 VKIDKEILKNLSVLFNCDKELTIS-----RLSPNRCPEKMFYLLMKTRNEHLSN 411  
 QY 331 STDFSSGCD-----GNSFYTOMGNETSNS 355  
 DB 412 SNFSFSSNDVDSARSLPRSTSVKTYTVDHATGEKHTYKLIQSSSIYSNLSLKSTS 471  
 QY 356 GGRVYIQDAEERP-----HSRYLRAYSSDRSGTSNSQOAKTYME 397  
 DB 472 AKGNVLSNTNPNTPKQPSASSSFNKKALHSHK--TQYASRSRNASRSLKSNSTOR 529  
 QY 398 RCHSAEMLSVKR-----RSGGGENEER-----YSTDNNANFN--PK 434  
 DB 530 NGNNASVTSVKNKIPEITGATVLOPIPSMAMNRGDEQONTKKNLTJTFGKSLINPOLIC 589  
 QY 435 EKTSSSGSFERRPDNNQALSNHLCPGKTPFPFADPPQETVYQ--WF-----480  
 DB 590 EAVFENDKENSKEPVSKTPVSO--LPPPPPIETPTSRNSVYKGTWLSLARERELAQ 647  
 QY 481 -----GNLQINAHLRKTEYD-----SISPNRDFQGHPI--QKDTSKNAWTDTV 524  
 DB 648 VQRMBARENKUAEELARKELEQEKRLAEKKEKRLQOGERELDEKQKLEKQKALERL 707  
 QY 525 KGNSDASD-----NAHSVQOQNTKMTALHAKREIILQOECVPSDPLSEBS 571  
 DB 708 QKQSHADEGLFASNRRSVTDMAPSSGMSLDPRAHWVSRANTIGSPNLSSS 761

RESULT 13  
 US-10-026-021-5  
 Sequence 5; Application US/10026021  
 Publication No. US20030027756A1  
 GENERAL INFORMATION:  
 APPLICANT: Hitooshi, Yasumichi  
 APPLICANT: Susan  
 APPLICANT: Jenkins, Yonchu  
 APPLICANT: Rigel Pharmaceuticals, Inc.  
 TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
 FILE REFERENCE: 021044-001210US  
 CURRENT APPLICATION NUMBER: US/10/026, 021

CURRENT FILING DATE: 2002-06-25  
 PRIOR APPLICATION NUMBER: US 60/309,632  
 PRIOR FILING DATE: 2001-08-01  
 NUMBER OF SEQ ID NOS: 8  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO.5  
 LENGTH: 400  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: DOMAIN  
 LOCATION: (1)..(400)  
 OTHER INFORMATION: human SNK mitotic kinase kinase domain  
 US-10-026-021-5

Query Match 11.0%; Score 560.5; DB 14; Length 400;  
 Best Local Similarity 41.6%; Pred. No. 4.6e-30;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNILGKSPAGYVYRAESIHITGLEVAIKMIDKAMKAGVQVQNEVKIKHCOLKPSILE 74  
 DB 85 GYVLGKGFPAKYEMTDLTNNKYVAKTIPIHSRVAKPHQRKIDKEIELRIHHGVVQ 144  
 QY 75 LNYFEDSNVYVLVLEMCHNGEMNRYLKNRVKPFSENEARHFMHQLITGMLYHSHGILH 134  
 DB 145 FYHYFEDKENIYILLECYGSRSMAILKAR-KVLTPEVRYVLRQIVSGIKYHLEQELIH 203  
 QY 135 RDLTSLNLLTTRNMTKIDFGLATOLKMPHEKHYTLCTPNYISPEIATRSAGLESVD 194  
 DB 204 RDLKGNFPIINAMELVKQVDFGLARLEPLHRRRTCTPNYLSVVLNKGHGESDI 263  
 QY 195 WSIGCMFYTLIGRPFPDPTVKNLANKVVLADYEMPSFLIEAKDLIHLRLRNPAURL 254  
 DB 264 WALGCVMTMLGRPFETITNLKETRYCRIRARYTMPSLSLPAKGLIASMLSKPEDRP 323  
 QY 255 SLSSVLDHPP 264  
 DB 324 SLDDIIRHDF 333

RESULT 14  
 US-09-771-161A-249  
 Sequence 249; Application US/09771161A  
 Patent No. US20020110811A1  
 GENERAL INFORMATION:  
 APPLICANT: LEVINE, et al.  
 TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
 FILE REFERENCE: 802620-2005.1  
 CURRENT APPLICATION NUMBER: US/09/771,161A  
 CURRENT FILING DATE: 2001-01-26  
 PRIOR APPLICATION NUMBER: 09/724,676  
 PRIOR FILING DATE: 2000-11-28  
 PRIOR APPLICATION NUMBER: 136776  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: 135619  
 PRIOR FILING DATE: 2000-04-12  
 NUMBER OF SEQ ID NOS: 273  
 SOFTWARE: Patentin version 3.0  
 SEQ ID NO 249  
 LENGTH: 685  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-771-161A-249

Query Match 11.0%; Score 560.5; DB 9; Length 685;  
 Best Local Similarity 41.6%; Pred. No. 9.7e-30;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNILGKSPAGYVYRAESIHITGLEVAIKMIDKAMKAGVQVQNEVKIKHCOLKPSILE 74  
 DB 85 GYVLGKGFPAKYEMTDLTNNKYVAKTIPIHSRVAKPHQRKIDKEIELRIHHGVVQ 144  
 QY 75 LNYFEDSNVYVLVLEMCHNGEMNRYLKNRVKPFSENEARHFMHQLITGMLYHSHGILH 134

```
Db      145 FHYHEDKENIYIILEYCSRRSMHILKAR-KVLTPEPVRYYLRQIVSGIKYIHQEIILH 203
QY      135 RDLTSLNLLITNNMIKIDFGIATQOLKMPHEKHYTLGTPNYISPEIATRSAGHLES DY 194
Db      204 RDLKLGNEFINAMELKVGDFGLARLEPLEHRRRTICGTPNYLSPEVLNKGHGCSEDI 263
QY      195 WSLGCMFYTLILGRPPFDITVKNTLKVVLADYEMPSFLSIEAKDLIHQLLRNPADRL 254
Db      264 WALGCVMTYMLLGRPPFETTNKETVRCIRBARYTMPSSLAPAKHLIASMTLSKNPEDRP 323
QY      255 SLSSVLDHPF 264
Db      324 SLDDIIRHDF 333
```

```
RESULT 15
US-09-771-161A-250
; Sequence 250, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 250
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-250
```

```
Query Match      11.0%; Score 560.5; DB 9; Length 685;
Best Local Similarity 41.6%; Pred. No. 9,7e-30;
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;
```

```
QY      15 GNLLGKSPAGVYRABSIHTGLEVAIKMIDKAMYKAGVORVONEVKIHCQLKPSILE 74
Db      85 GNVLRKGGPAKCYEMTDLTNNKVYAAKIIPHSRVAKPHOREKIDKEIHLRIHLHKAVQ 144
QY      75 LNYVEDSNVYVLVLEMCNEMNRYLNKRVKPFSENEARHFMHQTITGMILYHSHGILH 134
Db      145 FHYHEDKENIYIILEYCSRRSMHILKAR-KVLTPEPVRYYLRQIVSGIKYIHQEIILH 203
QY      135 RDLTSLNLLITNNMIKIDFGIATQOLKMPHEKHYTLGTPNYISPEIATRSAGHLES DY 194
Db      204 RDLKLGNEFINAMELKVGDFGLARLEPLEHRRRTICGTPNYLSPEVLNKGHGCSEDI 263
QY      195 WSLGCMFYTLILGRPPFDITVKNTLKVVLADYEMPSFLSIEAKDLIHQLLRNPADRL 254
Db      264 WALGCVMTYMLLGRPPFETTNKETVRCIRBARYTMPSSLAPAKHLIASMTLSKNPEDRP 323
QY      255 SLSSVLDHPF 264
Db      324 SLDDIIRHDF 333
```

Search completed: September 23, 2004, 21:04:13  
Job time : 72 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:58:26 ; Search time 23 Seconds  
(without alignments)  
4056.773 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKIEDPKVGNLTKGK.....KLQCLSSILMFSPNTPNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3927.5	77.3	925	2	A55748	protein kinase (EC
2	1921.5	37.8	465	2	B55748	protein kinase (EC
3	593	11.7	631	2	A57286	probable serine/th
4	591.5	11.6	576	2	S22127	protein kinase pol
5	561.5	11.1	682	2	A44493	serum-inducible ki
6	554.5	10.9	407	2	S52343	p46Bg265 protein -
7	548	10.8	389	2	S52242	protein kinase (EC
8	545	10.7	603	2	A47545	protein kinase (EC
9	545	10.7	603	2	A54596	protein kinase - m
10	544.5	10.7	1246	2	G89287	protein H39E23.1 (
11	536	10.6	603	2	S34130	serine/threonine-s
12	527	10.4	461	2	T14822	probable serine/th
13	527	10.4	1192	2	T18611	probable serine/th
14	526.5	10.4	1518	2	S37928	probable purine nu
15	514.5	10.1	683	2	T38254	serine/threonine-s
16	511.5	10.1	887	2	T20941	hypothetical prote
17	504.5	9.9	403	2	JCS974	autora-related kin
18	501	9.9	648	2	T43137	polo-like kinase-1
19	500	9.8	547	2	S44841	KO6H7.1 protein -
20	498	9.8	615	2	T29223	hypothetical prote
21	496	9.8	512	1	JC1446	serine/threonine-s
22	493.5	9.7	1142	2	S59359	GIN4 protein - yea
23	491	9.7	395	2	JCS975	autora-related kin
24	480	9.6	705	2	A48144	protein kinase CDC
25	489	9.6	441	2	B85362	hypothetical prote
26	487.5	9.6	384	2	T41298	ser/chr protein ki
27	486.5	9.6	502	2	T02306	probable protein k
28	486	9.6	915	2	S74283	probable protein k
29	484	9.5	512	2	T52633	serine/threonine-s

30	480.5	9.5	345	2	JC4665	protein kinase (EC
31	480	9.5	480	2	A86427	probable serine/th
32	479.5	9.4	504	2	T10449	probable serine/th
33	479	9.4	520	2	G86414	probable protein k
34	478.5	9.4	745	2	G01025	serine/threonine p
35	477.5	9.4	774	2	I48609	serine/threonine-s
36	476	9.4	445	2	T09903	probable serine/th
37	473	9.3	489	2	T04862	probable serine/th
38	470.5	9.3	456	2	T48478	serine/threonine p
39	469	9.2	367	2	S47923	probable protein k
40	469	9.2	453	2	G86141	protein T25K16.13
41	468	9.2	1081	2	S51899	probable protein k
42	465	9.2	592	2	T43402	probable protein k
43	464.5	9.1	473	1	S59941	serine/threonine-s
44	464	9.1	1558	2	T29253	hypothetical prote
45	462.5	9.1	633	1	A26030	serine/threonine-s

## ALIGNMENTS

## RESULT 1

A55748 protein kinase (EC 2.7.1.37) Sak-a - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Sep-1999

C/Accession: A55748

R/Node: C.; Motz, B.; Yousefi, S.; Heffernan, M.; Dennis, J.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 6388-6392, 1994

A/Title: Sak, a murine protein-serine/threonine kinase that is related to the Drosophila

A/Accession number: A55748; MUID:94294387; PMID:8022793

A/Accession: A55748

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-925 <F0D>

A/Cross-references: GB:129479; NID:548769; PIDN:AAC37648.1; PID:5487670

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C/Keywords: ATP; phosphotransferase

F/10-265/Domain: protein kinase homology <KIN>

F/18-26/Region: protein kinase ATP-binding motif

Query Match 77.3%; Score 3927.5; DB 2; Length 925;  
Best Local Similarity 78.6%; Pred. No. 3.8e-146;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;

QY	1	MATCGEKIEDPKVGNLTKGKSPAGYRAESHTGELVAIKMIDKAMTKAGVQRYONE	60
DB	1	MACTGERIEDPKVGNLTKGSPAGYRAESHTGELVAIKMIDKAMTKAGVQRYONE	60
QY	61	VKTHCOLKHPSTILELVNPFEDSNVYVLYLEMCHNGEMNRYLKNRVKPFSENEARHPMHQI	120
DB	61	VKTHCOLKHPSTILELVNPFEDSNVYVLYLEMCHNGEMNRYLKNRVKPFSENEARHPMHQI	120
QY	121	ITGMVLIHSHGILHRDLTISNLLITRNMMIKIADFGIATQMKPHKHYTLCTPYVISP	180
DB	121	ITGMVLIHSHGILHRDLTISNLLITRNMMIKIADFGIATQMKPHKHYTLCTPYVISP	180
QY	181	ELATRSAGLSDVMSLGCFFTLIGRPPEPTDTYKATLANKVVLADYEMPSFLSLEADQ	240
DB	181	ELATRSAGLSDVMSLGCFFTLIGRPPEPTDTYKATLANKVVLADYEMPSFLSLEADQ	240
QY	241	LHOLLRRNPADRLISVLDHPFMSRNSSTSKDGTVEDSIDGHATTSTLITSSST	300
DB	241	LHOLLRRNPADRLISVLDHPFMSRNSSTSKDGTVEDSIDGHATTSTLITSSST	300
QY	301	SISGSLFDRRLIGQPLPNKATVPEPKNSSTDFSSSGDGNSEFYQWGN--OETNSGRG	358
DB	301	SISGSLFDRRLIGQPLPNKATVPEPKNSSTDFSSSGDGNSEFYQWGN--OETNSGRG	358
QY	359	RVYQDAERPHSRVYLRRAVSDSGTNSQSQAKTYTMRCHSAEMLSVSKSGGGENEE	418
DB	359	RVYQDAERPHSRVYLRRAVSDSGTNSQSQAKTYTMRCHSAEMLSVSKSGGGENEE	418
QY	418	RVYQDAERPHSRVYLRRAVSDSGTNSQSQAKTYTMRCHSAEMLSVSKSGGGENEE	418
DB	418	RVYQDAERPHSRVYLRRAVSDSGTNSQSQAKTYTMRCHSAEMLSVSKSGGGENEE	418

```

Qy 419 RYSPPTDNNANIFNPFKEKTSSSSGSFERRPDNNQALSNHLCPGKTPFPFPADPTPOTETVQO 478
Db 411 -----LDENQHSNHCIGKTPFPFPADPTPQEMEVQO 442
Qy 479 MFGNQLMAHLKRTETYSISPNRDPQGHPLQKDTSNAMTDTKVKKNSDASDNASVK 538
Db 443 MFGNQLMAHLKRTETETHTVSPNRDPQDYPDLQ-DLLFNAMTDPRASKNADTSANVAVK 501
Qy 539 QONTMKTALHSKPEIIOQECVFSGDPLSEQSKTRGMERPWGONTLSITSPVLAHR 598
Db 502 QLSAMKTVSAHHNREVPORP--GLHHSQSKRSMESTLGYOKPLSITSPLLAHR 559
Qy 599 LKPIROKTKKAVSILDSEEVCEVLVREYASQEVYKVELQISSDGMTITTYYPNGGRGP 658
Db 560 LKPIROKTKKAVSILDSEEVCEVLVREYASQEVYKVELQISSDGMTITTYYPNGGRGP 619
Qy 659 LADPSPPTDNISSYSPNLPBKWKRYOVARFVQLVRSKSPKITTYFTYAKILMENS 718
Db 620 LADPSPPTDNISSYSPNLPBKWKRYOVARFVQLVRSKSPKITTYFTYAKILMENS 679
Qy 719 PGADPEWFTDGVKIKHKTEDFIQVLEKTKGKSYTLKSESVNSLKEEIKMTMDHANEGHR1 778
Db 680 PGADPEWFTDGVKIKHKTEDFIQVLEKTKGKSYTLKSESVNSLKEEIKMTMDHANEGHR1 739
Qy 779 CLAMESTIISBEERKTRAPPPPIIIGRKPGSTSPKALSPPSVDNYPTRDRASFNRWV 838
Db 740 CLAMESTIISBEERKTRAPPPPIIIGRKPGSTSPKALSPPSVDNYPTRDRASFNRWV 798
Qy 839 MHSNAPPTQAPILNPSWVTEGGLTTTASGTDISSSLDCLPKSAQLLKSPVNVGM 898
Db 799 VNSNAPPTQAPILNPSWVTEGGLTTTASGTDISSSLDCLPKSAQLLKSPVNVGM 853
Qy 899 ATQLTSGAVVQVRNDQSOLVVAQVSSISYSPNGQTRRGENEKLPDYIKQRLQCLSSI 958
Db 854 ATQLTSGAVVQVRNDQSOLVVAQVSSISYSPNGQTRRGENEKLPDYIKQRLQCLSSI 913
Qy 959 LLMFSNPTPNF 969
Db 914 LLMFSNPTPNF 924

```

```

RESULT 2
B55748
protein kinase (EC 2.7.1.37) Sak-b - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 19-Dec-1997
C/Accession: B55748
R/Code: C.; Mottu, B.; Yousefi, S.; Hefferman, M.; Dennis, J.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 6388-6392, 1994
A/Title: Sak, a murine protein-serine/threonine kinase that is related to the Drosophila
A/Reference number: A55748; MUID:94294387; PMID:8022793
A/Accession: B55748
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-465 <F0D>
A/Cross-references: GB:I29480
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP; phosphotransferase
F:10-265/Domain: protein kinase homology <KIN>
F:18-26/Region: protein kinase ATP-binding motif

```

```

Query Match 37.8%; Score 1921.5; DB 2; Length 465;
Best Local Similarity 84.5%; Pred. No. 2.5e-68;
Matches 370; Conservative 29; Mismatches 34; Indels 5; Gaps 4;
Qy 1 MATCHGKIEDEKVNLLGKSPAGVYRAESIHTGLEVAIKMIDKAMRYKAGVQVONE 60
Db 1 MAACIGKIEDEKVNLLGKSPAGVYRAESIHTGLEVAIKMIDKAMRYKAGVQVONE 60
Qy 61 VKIKHQLGHPISILLEYNFEDSNVYVLEMGCHGENMRVLAQNVKPESEAEAHFMHOI 120
Db 61 VKIKHQLGHPISILLEYNFEDSNVYVLEMGCHGENMRVLAQNVKPESEAEAHFMHOI 120

```

```

Qy 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPYISF 180
Db 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPYISF 180
Qy 181 EIAIRSAHGLSPVWSLGCMPFTLLIGRPPEPTDVKNTLVADYEMPSFLISIAKD 240
Db 181 EIAIRSAHGLSPVWSLGCMPFTLLIGRPPEPTDVKNTLVADYEMPSFLISIAKD 240
Qy 241 LTHQLLRNPADRLSLSVLDHPFMSRNSSTGSKDGLGYEDSIDSGHATISPAIRASST 300
Db 241 LTHQLLRNPADRLSLSVLDHPFMSRNSSTGSKDGLGYEDSIDSGHATISPAIRASST 300
Qy 301 SLSGSLPDKRLILGQPLPNKATVPKXKSTDPSSSDGNSFYTQGN--QETSNSGRG 358
Db 301 SLSGSLPDKRLILGQPLPNKATVPKXKSTDPSSSDGNSFYTQGN--QETSNSGRG 358
Qy 359 RYIODEERPHSRYYRRAVSSDSRGTNSOSQAKTYMTERCHSAMLVSQRSGGGENBE 418
Db 359 RYIODEERPHSRYYRRAVSSDSRGTNSOSQAKTYMTERCHSAMLVSQRSGGGENBE 418
Qy 419 RYSPPTDNNANIFNPFKEK 436
Db 418 RYSPPTKSNVNVLTSLNLT 435

```

```

RESULT 3
A57286
probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse
C/Species: Mus musculus (house mouse)
C/Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997
C/Accession: A57286
R/Donohue, P.J.; Alberts, G.F.; Guo, Y.; Winkles, J.A.
J. Biol. Chem. 270, 10351-10357, 1995
A/Title: Identification by targeted differential display of an immediate early gene encoc
A/Reference number: A57286; MUID:95247749; PMID:7730342
A/Accession: A57286
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-631 <DON>
A/Cross-references: GB:U21392; GB:U22434
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP; phosphotransferase
F:61-315/Domain: protein kinase homology <KIN>

```

```

Query Match 11.7%; Score 593; DB 2; Length 631;
Best Local Similarity 33.1%; Pred. No. 1.9e-16;
Matches 122; Conservative 64; Mismatches 111; Indels 72; Gaps 4;
Qy 15 GNLLGKSPAGVYRAESIHTGLEVAIKMIDKAMRYKAGVQVONVKTIKQKHPISIE 74
Db 66 GRLLGKGFARCEATDTESGIYAVVQVIRQSRVAKPHQREKILNELHRDLQHRHYR 125
Qy 75 LNVYFEDSNVYVLEMGCHGENMRVLAQNVKPESEAEAHFMHOIITGMLYHSHGILH 134
Db 126 FSHHPEADNIVYFLEICSRKSLAHYKWR-HTLLEPEVARYYLRQILSGLYHQRGIH 184
Qy 135 RDLTTLNLLTRNNMIKIADFGIATOLKMPHEKHYTLGCPNYISPEIRSAHGLSESV 194
Db 185 RDLKGNPFITDMELKQVDFGLAARLEPPEQKKTICGTENVVAEVLRLRGHGEADV 244
Qy 195 WSLGCMFTLLIRPPEPTDVKNTLVADYEMPSFLISIAKDILHQLRRNPADRL 254
Db 245 WSLGCMFTLLIRPPEPTDVKNTLVADYEMPSFLISIAKDILHQLRRNPADRL 254
Qy 255 SLSSVLDHPFMSR-----NSTSKDLG 277
Db 305 SLGILRHFPFTKGYPPDLVPVSSCYTVPDLTPPNPANSLEFAKVTKSLFGRRKNKXHS 364
Qy 278 TWEDSID-----SGHATISAITASSSTISGSLFDKRLRLILGQPLPNKATVPKXKSS 331
Db 365 EPQDNVSLAPVVGQAPASLLETAAEDSSPRGTL----- 399
Qy 332 TDFSSSGDG 340

```



```

Db      137 LEDPEIGRLGKGRGVNYLAREERESKTLALKTLPFSQLEKAQVEHQLRREVAIQSHLR 196
      69 HPSLLELVNYPEDSNVYVLVLEMGCHNGEMNRYLKNRYKVPSENEARRHFMQIITGMVLH 128
      197 HENILRLYGFYHDSRVYLLIDYAPGGLFRELQKCTR-FDDQSAWYIKQLAELALYCH 255
Qy      129 SHGILHRLDTLSNLLTRNNNIKTADFGLATQLMKPHKHYTLGCTPNYISPEIATRSAN 188
      256 SKKVIHRDIPKRNLLSGNELKTLADFGMS--VHAPSSRTLLCGTLDYLPPEMIEGRMH 313
Db      189 GLSDVWSLGCMTYLLIGRPEDPTVKTINKVNLADYEMPSFLSIEAKDLIHQLLR 248
      314 DETYDLMSLGYLCYEFVVGKPPFTDTHQETRYRISKVEFOYPYVSEBARDLVSKLLKH 373
Qy      249 NPADRLSLSSVLDHPFMSRNSSTKSD 275
      374 NPNHRLPLKGVLEHPIWIKNSQLKKKD 400

```

RESULT 7  
 S52242  
 Protein kinase (EC 2.7.1.-) p46Xlg22 - African clawed frog  
 C/Species: Xenopus laevis (African clawed frog)  
 C/Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
 C/Accession: S52242  
 R/Kogut, C.; Le Guellec, R.; Paris, J.; Coucuerier, A.; Philippe, M.  
 Submitted to the EMBL Data Library, October 1992  
 A/Description: Bg2, selected by differential screening encodes a new Xenopus protein kin  
 A/Reference number: S52242  
 A/Accession: S52242  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-389 <R0G>  
 A/Cross-references: EMBL:Z17206; NID:9609279; PIDN:CAA78914.1; PID:9609280  
 C/Superfamily: kinase-related transforming protein; protein kinase homology  
 C/Keywords: ATP; phosphotransferase  
 F:119-371/Domain: protein kinase homology <KIN>  
 F:127-135/Region: protein kinase ATP-binding motif

Query Match 10.8%; Score 548; DB 2; Length 389;  
 Best Local Similarity 39.9%; Pred. No. 6.5e-15;  
 Matches 107; Conservative 60; Mismatches 97; Indels 4; Gaps 3;

```

Qy      9 IEDFKVGNLKGSPAGYVRAASINTGLVAKIMDKAMVYAGVQVONEVKTHCOLK 68
      118 LEDPEIGRLGKGRGVNYLAREERESKTLALKTLPFSQLEKAQVEHQLRREVAIQSHLR 177
Db      69 HPSLLELVNYPEDSNVYVLVLEMGCHNGEMNRYLKNRYKVPSENEARRHFMQIITGMVLH 128
      178 HENILRLYGFYHDSRVYLLIDYAPGGLFRELQKCTR-FDDQSAWYIKQLAELALYCH 236
Qy      129 SHGILHRLDTLSNLLTRNNNIKTADFGLATQLMKPHKHYTLGCTPNYISPEIATRSAN 188
      237 SKKVIHRDIPKRNLLSGNELKTLADFGMS--VHAPSSRTLLCGTLDYLPPEMIEGRMH 294
Db      189 GLSDVWSLGCMTYLLIGRPEDPTVKTINKVNLADYEMPSFLSIEAKDLIHQLLR 248
      295 DEKVDLMSLGYLCYEFVVGKPPFTDTHQETRYRISKVEFOYPYVSEBARDLVSKLLKH 354
Qy      249 NPADRLSLSSVLDHPFMSRNS-STKSD 275
      355 NPNHRLPLKGVLEHPIWIKNSQLKKKD 382

```

RESULT 8  
 A47545  
 Protein kinase (EC 2.7.1.37) Plk - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1999  
 C/Accession: A47545  
 R/Clay, P.J.; McEwen, S.J.; Bettoncello, I.; Wilks, A.F.; Dunn, A.R.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 4882-4886, 1993

A/Title: Identification and cloning of a protein kinase-encoding mouse gene, Plk, related  
 A/Reference number: A47545; MUID:93281660; PMID:8099445  
 A/Accession: A47545  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-603 <CIA>  
 A/Cross-references: GB:L06144; NID:9309461; PIDN:AAA39948.1; PID:9309462  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C/Keywords: ATP; phosphotransferase  
 F:51-305/Domain: protein kinase homology <KIN>

Query Match 10.7%; Score 545; DB 2; Length 603;  
 Best Local Similarity 27.4%; Pred. No. 1.3e-14;  
 Matches 153; Conservative 93; Mismatches 192; Indels 120; Gaps 20;

```

Qy      15 GNILKGSPAGYVRAASINTGLVAKIMDKAMVYAGVQVONEVKTHCOLKPSILE 74
      56 GRPLGKGFPAKCFEISDADYKEVFAKQIVPSLLKPKQKEMSWEISIHSLAHQHYVG 115
Db      75 LNYPEDSNVYVLVLEMGCHNGEMNRYLKNRYKVPSENEARRHFMQIITGMVLHSHGILH 134
      116 FDDPEFSDPFVYVLELCRRRSLLLELHRR-KALTEPARAYLRQIVLGCQYLHNGYIH 174
Qy      135 RDLTTSNLLTRNNNIKTADFGLATQLMKPHKHYTLGCTPNYISPEIATRSANGLSESDV 194
      175 RDLKGNLFNEDLEVKIGDGLATKVEYEGERRKTLGCTPNYLAPEVLSKGGHSEVDV 234
Qy      195 WSLGCMFTLLIGRPEDPTVKTINKVNLADYEMPSFLSIEAKDLIHQLLRNPADRL 254
      235 WSGICIMVTLVVGKPPFTSCIKETVYLKKNQEVSPGHNPVAAASLIQKMLQYDPTARP 294
Db      255 SLSSVLDHPFMSRNSSTKSDLGTVEDSIDSGHATIS--FAIRASSTISGSLFD-K 309
      295 TIHELINDEFT-----SGVLPARLPICLTLPREFSIAPSSLDPS 336
Qy      310 RRL-----IGQPLPKMTVPFNKKSSTDPSSSGDG-----NSFYQWNGQETSN-SGR 357
      337 RKPLKLVINKGVENPLPDR-----PREKEPVAREINMAIECHLSLDLQQLTSVNASKSESR 392
Qy      358 GRVQDAEERH-----SRVLRAYSSDRSCTS-----NSQSQATYTWER 398
      393 GLVQOELEEDACPIPIFWSSKMD--YSDKYGLGYOLCNSVGVLPNDSTRLILY---- 445
Db      399 CHSAEMLSVSRSGGSENERYSPTDNAN-----IFNPFKKTSSSSGSFRRPNNG 451
      446 -NDDDSLQYIERD-----TESYLVSSHPSLSMKKTILAYFR-----N 484
Qy      452 ALSNHL-C-PGKTPP-PADPTPQTEVQWF-----GNTQIN-----A 487
      485 YMSHLLKAGRNITPRRGDEIARLPYLRFTWFRRTSAIILHLSNGTVQINFPQHTKILC 544
Qy      488 HLKRTYDSISPRDPQ 505
      545 PLMAAVTY--INEXKDFQ 560

```

RESULT 9  
 A54596  
 protein kinase - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999  
 C/Accession: A54596  
 C/Lake, R.J.; Jelinek, W.R.  
 Mol. Cell. Biol. 13, 7793-7801, 1993  
 A/Title: Cell cycle- and terminal differentiation-associated regulation of the mouse mRN  
 A/Reference number: A54596; NUID:94067140; PMID:7902533  
 A/Accession: A54596  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: mRNA  
 A/Residues: 1-603 <RBS>  
 A/Cross-references: GB:L19558; NID:9403473; PIDN:AAA16071.1; PID:9403474  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 F:51-305/Domain: protein kinase homology <KIN>

Query Match 10.7%; Score 545; DB 2; Length 603;  
Best Local Similarity 27.2%; Pred. No. 1,3e-14;  
Matches 152; Conservative 95; Mismatches 190; Indels 122; Gaps 20;

QY 15 GNLLGKSPAGVYRAISHTGLEVAIKMIDKAMVYQAVQVQNEVKIHCQKPSILE 74  
DB 56 GRLLGKSPAGVYRAISHTGLEVAIKMIDKAMVYQAVQVQNEVKIHCQKPSILE 115  
QY 75 LNYVPEDSNVVYLVLEMCHNGEMNRVLYKRVKPFSENEAHPMHQITGLYLHSHGILH 134  
DB 116 FHDFFSDSPFVFNVLCCRRLSLLEHKKR-KALTEBAKYLQVLGCGYLRHQVH 174  
QY 135 RDLTSLNLLTRMNIKIADFGLATOLKMPHEKHYTLGTPNYISPEIATRSASHGESPD 194  
DB 175 RDLKAGNLLPLNELBKVIGDFGLATVYKYGKRTILGTPNYIABEVLKSHSEVDV 234  
QY 195 WSLGCMFYTLIGRPPEPDTVTQNTLKVADLYEMPSPFLSIKADLIHQLLRNPAERL 254  
DB 235 WSLGCMFYTLIGRPPEPDTVTQNTLKVADLYEMPSPFLSIKADLIHQLLRNPAERL 254  
QY 255 SLSSVLDHPMSNLSKSDLTVEDSIDSGHATIS-TAITASSSTISGSLPD-K 309  
DB 295 TTHELINDERFT-----SGYIPARLPITCLTIPRPSIAPSSLDPS 336  
QY 310 RRL-----IGQLPNKQTVFPKNSSTDPSSSGD-----NSFTQMGNETSN-SGR 357  
DB 337 RRLKPLKLVNGVFNPLDR-----PREKEPVRVRETNEAIECHSLDLQOLTVNASKPESR 392  
QY 358 GRVYIDAEERPH-----SRVLRAYSSDRSGTS-----NSQSAKTYTMR 398  
DB 393 GLVROGEADPACIPFVWSKWD-----YSDKYGIGVQLCONSGVLFNDSTRILV----- 445  
QY 399 CHSAEMLSTSKSGGSENERISPTDNAN-----IPVPEKETS-----SGSERPD 448  
DB 446 -NDGSLQYIERG-----TESYLVSSHPSMLKITTLLNFPNYSHEHLKAGANITPR 500  
QY 449 NNQALSNHLCPGTPPEPDPPTQTEVQWF-----GNLQIN----- 486  
DB 501 EGDELA-----RLPY-----LRWFRTRSAIILHLSNGVQLNFPDHTKLL 543  
QY 487 AHLKRTEDYSISPNRDFQ 505  
DB 544 CPLMAVTV--INERKDFQ 560

RESULT 10  
G89287  
protein H39E23.1 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: G89287  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MIMD:9906963; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G89287  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1246 <STO>  
A:Cross-references: GB:chr\_V; PIDN:CAM09532.1; PID:g3878100; GSPDB:GN00023; CESP:H39E23.  
C:Genetic:  
A:Gene: H39E23.1  
A:Map position: 5

Query Match 10.7%; Score 544.5; DB 2; Length 1246;  
Best Local Similarity 22.8%; Pred. No. 3e-14;  
Matches 252; Conservative 167; Mismatches 436; Indels 249; Gaps 38;

QY 12 FKVGNLLGKSPAGVYRAISHTGLEVAIKMIDKAMVYQAVQVQNEVKIHCQKPS 71  
DB 116 VSLGCMFYTLIGRPPEPDTVTQNTLKVADLYEMPSPFLSIKADLIHQLLRNPAERL 254

DB 60 YKLLTKIGKNFAPKVLAKHVTGHEVAIKIDKALNPSL-QKLFREVKIMQDLNHN 118  
QY 72 ILELVNPEDSNVVYLVLEMCHNGEMNRVLY-KNRYKPFSENEAHPMHQITGLYLHLS 129  
DB 119 IYKLVNPEDSNVVYLVLEMCHNGEMNRVLY-KNRYKPFSENEAHPMHQITGLYLHLS 175  
QY 130 HGLHBDLTLNLLTRMNIKIADFGLATOLKMPHEKHYTLGTPNYISPEIATRSASH 188  
DB 176 KNIHRLKAEVNLDDQDNITKIADFGSNTSISL-KLDTGCGSPVAPAEIPLSGKVD 234  
QY 189 GLESDVWLSGCMFYTLIGRPPEPDTVTQNTLKVADLYEMPSPFLSIKADLIHQLLR 248  
DB 235 GEVDVWLSGCMFYTLIGRPPEPDTVTQNTLKVADLYEMPSPFLSIKADLIHQLLR 294  
QY 249 NPADRSLSSVLDHPMSN----- 268  
DB 295 NPDARSLSSVLDHPMSN----- 354  
QY 269 --SSTKSD-----LGTVEDSIDGHATISAITASSSTISGSLPDKRRLIGQPLP 319  
DB 355 SVEKEKFEEDIHATYLLGERSKDMASEITMAQSLSHSINVSSL-----GQHPA 406  
QY 320 NKAT--VPPKNSSTDFSSGSDGNSFYQMGNETSNGRGVYIDAEERPHSKYLRPA 376  
DB 407 GVTREHVTSSASGSSASPSRYSASATATASITAGSA--LASNAVAOKH-----QOS 459  
QY 377 YSDRSGTSNOSQAKTYTMRCHSAEMLSVKRSKSGGSENER-----YS 421  
DB 460 SAAPSSGSSSSSRSSQNDAAATAGTVMGSTRGGVMAQPPSRQATISLLOPPSK 519  
QY 422 PTDNNANIN-----PFKEKTSSSSGSFER-----PDNNQALSNHLCP-- 459  
DB 520 PNTNTQIQAQIPPLFRNSTATSSAAQPSGTITGRKIADPKGRIPLNSTAVGCHTANG 579  
QY 460 -----GKTPPEPADPTEVQWFGLQNLNHLKRT-----TEYDISPNRDFQHP 508  
DB 580 AVALANNQIP-SHRDAQOOQYNNQOLSTTMSKLNKTPPAAGTATATSSSSSASTSA 638  
QY 509 DLQKOTSXNAWTDYKXKSDASDNASHVKQONTMKYMAL--HSKP--EIIQOECVFGS 564  
DB 639 PLOKSGSQLSHAPEVIREDDEN-NSENQGNVPLIGVGQBPOTSBAVQVPEDTSS 697  
QY 565 DPLSEQSKTRGMBP-----PWGQNTLRISITPLVAHRLKPIROKTKKAVSLISEEVC 620  
DB 698 DKEQOQKASSETPKESNPIVQNLHLNLSLSDSSAATSEYFTRRRGIGARRSEPA 757  
QY 621 VELVEYASQEVYKVLQISSD-----GNTI-----TIYPNGRGFPPLADRPSPPT 667  
DB 758 ATPRRHQTMVVDARHLQPPDTPRPHEDTTLDRQMALVYSTASSHTRGVLPPTPT 817  
QY 668 DNISKSPFNLPEKWKRYKQVARSFQVYRSKSPKT-----YFTRVAK----- 711  
DB 818 SNSTSSSFVEPLTH-----VAASPDITTTTPKSVTVTSYPRRRPSPFM 864  
QY 712 ----CIENSPGADPEVYDGVKIHKTEDTQVLEKTKGSKYTLKSESVNSLKEIRH 767  
DB 865 VLVLLICNG-----LRLM--PSMTHQSPM-----PQGMATMSESLTL 903  
QY 768 YMDHANEGRICALESIIEBERKTRAPFPPIIIGRK-----PGST 810  
DB 904 SESGQGTGPTVATGPPQATSGQMSRSATYTSANMAGSGGAAATAATNOLGAPSSPT 963  
QY 811 SS-----PKALSPSPVSNVPTDRAPFNMWHSASAPFOAPLNSMTNBEGLGT 864  
DB 964 GASSQGYHRAKAPSSSSSTNPPHQQLTN-----ASPSVPSYQIP--TSTAVNVT 1015  
QY 865 TTASGTDISSNLKDLCKSAQLKSVFYKNGVMAATQLSGAVWVQFNDGSLVQA-CV 923  
DB 1016 STGMPTSSSSSA-----FPNTANROTFFHK-----TEKXKG-----DSSDEIGTTPEN 1061  
QY 924 SSISYTPS--NGQTRVGENEKL 944  
DB 1062 VSGATGATPANNNAEATWISKSL 1085



```

A:Accession: T18611
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W11.1>
A:Cross-references: EMBL:281027; PIDN: CAB54179.1; GSPDB: GN00023; CESP: H39E23.1a
A:Experimental source: clone AH10
A:Accession: T18610
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W11.2>
A:Cross-references: EMBL:281027; PIDN: CAB54178.1; GSPDB: GN00023; CESP: H39E23.1b
A:Experimental source: clone AH10
R:McMurray, A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z19696
A:Accession: T23144
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W11.3>
A:Cross-references: EMBL:296102; PIDN: CAB54263.1; GSPDB: GN00023; CESP: H39E23.1a
A:Experimental source: clone H39E23
A:Accession: T23143
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W11.4>
A:Cross-references: EMBL:296102; PIDN: CAB54262.1; GSPDB: GN00023; CESP: H39E23.1b
A:Experimental source: clone H39E23
C:Genetics:
A:Gene: CESP:H39E23.1a; CESP:H39E23.1b
A:Map position: 5
A:Insertions: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 535/3; 631/1; 825/2; 914/3; 992
C:Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific pro
F:1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #8
F:1-487, 536-1192/Product: probable serine/threonine-specific protein kinase, short splice
Query Match 10.4%; Score 527; DB 2; Length 1192;
Best Local Similarity 25.8%; Pred. No. 1,4e-13;
Matches 183; Conservative 108; Mismatches 252; Indels 166; Gaps 20;
QY 12 PAVGNLKGSGPAVYRAESHTGLVAIKMDKAMTKAGVQVQVQVYKHCQLKPHS 71
DB 170 YKLLKTIKGKGVAKVLAHGVHVAIKIIDKTLNPSSTL-QKLFREYKIMKQLDHNP 228
QY 72 ILELVYFEDSNVYVLYLEMCHNGENRNYL--KNRKPSSENAHFMQIITGMVLYLS 129
DB 229 IYKLVQVMTBOTLYLVLYASGGEVFDYLVAHGRMK---EKARAKPRQIVSAVGYLS 285
QY 130 HGLIHRDLTSLNLLTRNNNIKIADFGLATQLMKPEKHYTLGTPNYISPEI-ATRSAAH 188
DB 286 KNIIRHDLKAEVLLDQDNKIKADGFSNTFSLGN-KLDTFGSGFPYAAPELFSKKYD 344
QY 189 GLESDVWSIGCMFYTLIRPPPTDTVKNLTKVVLADYEMPSFLISEAKDLIHQLLR 248
DB 345 GEVVDVWSIGVILYTLVSGSLPFDGQNLKELRERVRGKRIIPFYMSTDCENLKKFLVY 404
QY 249 NPADRLSSVLDHPMSN----- 268
DB 405 NQORSSLDNIMKDRMMNNGYEDDELKPIEPKQIDQRIEKLQIFQLGPNKALILE 464
QY 269 --SSTKSKD-----LGTVEDSIDSGHATISTATPASTSISGSLPDKRRLILGQPLP 319
DB 465 SVKEKFEFEDIHATYLLGGRKSDMDASEITMAQSLSHSSINVSSSL-----GQHA 516
QY 320 NKMT--VPKPKASSTDFSSGSGNSFYQWGNQETNSNGRGVYIDAEERPHRYLRA 376
DB 517 GVITREHYVSSASGSSASPSRYSRSATATGASITAGSA---LASAANAQKH---QOS 569
QY 377 YSDRSGTNSGQATTYTMRCHSAEMLSVSKRSGGSENER-----YS 421
DB 570 SAAPSSGSSSRKSSQNDAAATAGGTVAWSGTRHGQVQWRAOPTSRQATISLLQPPSYK 629
QY 422 PTDNNANINFN---PFKEKTSSSSGSFER-----PDNNQALSNHLCP-- 459

```

```

DB 630 PSSNTTQIAQIPELPFNBNSTATSSAAQSTGILTRKLIADPKGRILPINSTAVQGHRTATG 669
QY 460 -----GKTPPPADPTQETETVQWFGNLIQINAHLRKT-----TEYDSISPRDQGH 508
DB 690 AYAANNNGGIP-SHRDAQOOQVYNQLTSTSMKSLINKPKAPAGCTATATSSSSSASTST 748
QY 509 DLQKDTSK--NMAVTDKVKKN-----SDASDNH 535
DB 749 PLQSGSSQISHAFTEVIREDDENNENGNVPLIGVGPTSPAVQPTEDATSSSD 808
QY 536 SVKQNTMYKMTALHSKPEIIQOE-----CVFGSDPLSEQSTRTG 575
DB 809 KEQOOQKASSETPKSKPSMIMHSPMPSPQMTAMESLKSSESGQTGG 857

RESULT 14
S37928
probable purine nucleotide-binding protein YKJ101w - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKJ453
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C:Accession: S37928; S39084
R:Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzos
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37920
A:Accession: S37928
A:Molecule type: DNA
A:Residues: 1-1518 <CHR>
A:Cross-references: EMBL:Z28101; NID:g486168; PIDN: CAB81941.1; PID:g486169; MIPS: YKJ101w
A:Experimental source: strain S288C
R:Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara
Yeast 9, 1149-1155, 1993
A:Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loca
protein kinase.
A:Reference number: S39084; MUID:94078677; PMID:8256524
A:Accession: S39084
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1518 <PAL>
A:Cross-references: EMBL:X71133; NID:g431205; PIDN: CAB50456.1; PID:g431215
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:HSL1
A:Cross-references: SGD:S0001584; MIPS: YKJ101w
A:Map position: 11L
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein kin
F:79-369/Domain: protein kinase homology <KIN>
F:79-86/Region: nucleotide-binding motif A (P-loop)
F:87-95/Region: protein kinase ATP-binding motif
F:85/Binding site: ATP/GTP (lys) #status predicted
Query Match 10.4%; Score 526.5; DB 2; Length 1518;
Best Local Similarity 22.3%; Pred. No. 1.9e-13;
Matches 253; Conservative 182; Mismatches 429; Indels 271; Gaps 42;
QY 7 EKIEDPKVGNLKGSGPAVYRAESHTGLVAIKMDKKA-----MYKAGN 53
DB 76 DTVGPKKGLKTYLKGSGGVRLAKNMETQOLAIVPKKKAFFHCSNNGTPVNSYSSSS 135
QY 54 VOR-----VQNEVKIKHQLKHPSSILELVYFEDSNVYVLYLEM 91
DB 136 VTSNVSPSIAAREHNSHOTNPYGERIRIVMKLISHNNVALLPFWENKSELVLYLE 195
QY 92 CHNGEMRYLKORVXPFSENEAHFMHQIITGMVLYSHGILHRDITLSNLLI-TRNNNI 150
DB 196 VQGGELFDYLVSKGK-LPREAIIHYKQIVGVSCHSFNICHRLDKPENLLLDKKGRRI 254
QY 151 KTADPGLATQLMKPHKHYTLGTPYISPELAT--RSAGLESDVWSIGCMFYTLILGR 209
DB 255 KIADFGMAA-LLEIPNKLTKTSCSPHYASPELVMGVPGGSDVWSCIVLFAALTGH 313
QY 210 PFDTDTVKNLTKVVLADYEMPSFLISEAKDLIHQLLRNPADRLSLSSVLDHPF----- 264

```

```

Db 314 PFNDNNTKLLKLVGSGKQMPNLSSEARDLSKILVLDPEKITTOELIKPFLKTD 373
QY 265 -----MSNNSSTKSD-----LGVESIDSGHA--TISALTAASST---SIS 303
Db 374 DLPEVKVLRMRKDNNAFGKSNBDLHLNNVSPSIVTLHSKGRIDESIIRSLQIIMHGVS 433
QY 304 GSLFDKRLILIGPLFNKMTVPF-----KVKSTDPSSGDDGNSFYTONGNQTNSGKG 358
Db 434 RELITAK--LQKPMEEKLFYLLQYKORHSISLSSSENKSKAT-----478
QY 359 RVIQDAEERPHSRYLRAVSDSGTSNSQAKTYTMRCHSAEMLSVSKSG-GENE 417
Db 479 ---ESSVNERLEYSK--TANNTGSRSENNVKTLSLHSESDTSTVQNNALITGVAT 533
QY 418 ERYSPTDNNANIFNFKEKTSSSSGSFEPDNNQALSNHL-CPGKTPPPAD-----468
Db 534 EIMAVY-----LAQSQFSINTLSQPSDKAAEAVTLPPALPIFNASSRIFRNS 584
QY 469 ----PYQTEVQQWFGNQLINAHLRKTEYDISISPNRDFOGHPDLQKDTSKNMTDTKV 524
Db 585 YNISRSRRSLRLSRLSLSTSLSTRETYHDNEMF-----LPQLPKSPKYSLRRAI 638
QY 525 KKNSDASNAHVYKQONTKMTALHSKPEIILQECVFGSDPLSEQSKTRGKPEPWGYON 584
Db 639 HASPSTKSIHKSIRKNIATVAAARTLQNSAKRSLSYLSQISKRK-----LNL 688
QY 585 RFLRSTISPLVAHRLKPIROKTKKANVSLDS-----EECVELY-----624
Db 689 NDLPVFPDPLPSK--KPASENVNKSSEPHLESQDELLCDQIILFNALDRLLEEEDNE 746
QY 625 KEVASQEVYKAVLQISDQNTI---TIYYPNGRGFPPLA-----DRPSPDTNISRYSP 675
Db 747 KERDTQROQNDTKSSADFTTISGVSTKENGEPEPTIKENQFMYSKPEBNSGLSS 806
QY 676 DNLPEYTKRKQYASRFVQLVASKPKITYFTRYAKCIEMENS PGADFVWFYDGYKTHK 735
Db 807 FPIFEK---ENTLSSSYLEOKPKRALMSDITNSFNKNKQ-----EGMRLEK 851
QY 736 TEDFTQVLEKTKSYTLK---SESEVNSLKEE---IKMYMD--HANEGRICLAEISIT 787
Db 853 KIQROQLQKNDPRSPLEKIQHOELRVNSLPMDQKPSISLDPRRNVIQSPVNSKVESLQ 911
QY 788 EEEKTRSAFPFPIIIG-----RKPGSTSPKALSPPSVDSNY--PTRDRASFNR 836
Db 912 GLKFKKEPASHWTHREGSLFMSEHVEDEKPVKASD-----VSLESSVPLTVAATSR 964
QY 837 WMMHGAAPT-QAPF-----LNPMTNNEGILGTTASGTD-----871
Db 965 DSVLAESSTIQKPMLSLPSFPLNTSM--TFKNLSQIILADDDGDKHLSVPQNSRVSAMSH 1023
QY 872 -ISSNSLKDCPKSAQLLKSVFK-----NVGMATQL-----TSGAV 907
Db 1024 PLRKQSAKISLTPRNNANLNSVKKNQSGPSYLSNDLGDISDMTFAMEIPNTTTAQAI 1083
QY 908 WYQFND-----GSQLVVQAGVSSISYTS--PNCQTRTYGENEKLPLY 947
Db 1084 QLMNNDTNKKNINTSPKASSFTKEKVIKSAAYISKKEKPDNDTNY-----IPDY 1133

```

## RESULT 15

```

serine/threonine-specific protein kinase p101 (BC 2.7.1.-) - fission yeast (Schizosaccha
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: J38254; T45128
R:Brown, D.; Churcher, C.M.; Bartrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z21781
A:Accession: J38254
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-683 <BRO>

```

```

A:Cross-references: EMBL:Z98559; PIDN:CA811167.1; GSPDB:GN00066; SPDB:SPAC23C11.16
A:Experimental source: strain 972h-; cosmid c23C11
R:Ohkura, H.; Hagan, I.M.; Glover, D.M.
Genes Dev. 9, 1059-1073, 1995
A:Title: The conserved Schizosaccharomyces pombe kinase p101, required to form a bipolar
A:Reference number: Z22921; MUID:955262899; PMID:7744248
A:Accession: T45128
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-683 <OHK>
A:Cross-references: EMBL:X85758; MID:9887640; PIDN:CA59766.1; PID:9887641
A:Experimental source: strain 972 derivative
C:Genetics:
A:Gene: p101; SPAC23C11.16
C:Function:
A:Description: required to form a bipolar spindle and early in the regulatory cascade the
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase; protein kinase

Query Match 10.1%; Score 514.5; DB 2; Length 683;
Best Local Similarity 24.4%; Pred. No. 2.4e-13;
Matches 176; Conservative 106; Mismatches 247; Indels 191; Gaps 22;

QY 18 LGKSPAGVYRAESIHTEGLVAIKMIDKKAMVQVQVONEVKIHCOLKHPSTLELYN 77
Db 47 IGGGFACCFKVKD--NYGNITAAKVIKRSLONDKTKLKGELIKQSMHSHPIVGTID 105
QY 78 YFEDSNVYVLEMGCHNGENNRVYLNKRVKPPSENEARHFMQIITGMULYHSHGLHRDL 137
Db 106 CFEDSTNIYILILECEHKSIMELLKR--KQLEPEVRVIMQIIGALCYMKHKKVYIHRDL 164
QY 138 TLSNILLTRNNIKIADRGATOLKMPHEKHYTLGCPNTYNSPIA--TRAHGLESVW 195
Db 165 KLGNIIMDESNNVAKIGDFGLALLMDDEERMTTCGPNYIAPEILFNSKSGHFEVDLW 224
QY 196 SLGCMFYTLIGRPFPDVTYKNTLNKVLADYEMPS--PLSIAKDLIHOLLRRNPADR 253
Db 225 SAGVVMYALLIGKPPQDKEVKTIRKIKANSYSPFSNVDSIAKXOLISLTHDPBIR 284
QY 254 LSLSSVLDPHPSNRNNSSTKSKDLGTVEDSIDSHATTSTAITASSSTISGSLPDKRL 313
Db 285 PSIDIDVDFP-----PHTGYMAST-----304
QY 314 IGQPLPMKMTVPFKKST-----DSSSGDGNSTYQNG-----NOETS-- 354
Db 305 LPDELHSMPIWPSQSKSSFQRLNDFVASGVGFNSAGVEKXKPYALTDEVDNDRI 364
QY 355 -----SGRGVYQ---DAERPHSRYLRAVSDSGTSNS-----QSOA--- 391
Db 365 LPVSLPDRVNPVWKIGPETKVPKSLSTALHAARSTOGLSRVYVLRSESPQVPT 424
QY 392 KTYMERCHSAEMLSVSKRSGGGENERYSPTDNNANIFNFKEKTSSSSGSFERRPDNQ 451
Db 425 KSAVTEQVEPIQLI---RSLSANVTASRLSKVGMKDKDIWISVYKTKLKG- 471
QY 452 ALSNHLCPGKTPPPADPTQTEVQQWFGNQLINAHLRKTEYDIS-----499
Db 472 -----MALEAHTHALTSDEADSEPVLPFTTKVNDY 500
QY 500 PNRDFOGHPDLQKDTSKNMTDTKYKKNSDASNAHVSKQONT-----MKYMTALH 550
Db 501 SNKYGGLGVLSDSEVGHPNDTSLPSADBEVVEYALHPDTEIKPIYIPASVPSIR 560
QY 551 SKPFIQGEVFGSDPLSEQSKTRGKPEPWGYONRTLRSLTSPVVARLKPFRKT-----606
Db 561 SKIQLRHFYSWQNSKAVQDSESFEPK--KNST--SVYMLFMQHYLR--TROALMFR 614
QY 607 -----KAAVS-----LIDSEVCELVKAYASQ-----EVKVEVLO 638
Db 615 SNGIFQNPADHRKAVVISTARKIIVLDKREKREVELPQZASAFSEDLRSRLKTIIRBTL 674

```

Mon Sep 27 12:38:15 2004

Search completed: September 23, 2004, 21:02:30  
Job time : 27 secs

us-10-026-021-2.rpr

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 23, 2004, 20:59:21 ; Search time 22 Seconds  
(without alignments)  
2276.236 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078  
Sequence: 1 MATCGEKEDPKVGNLJGK.....KQCLSSILMFSPNPNH 970

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgnt2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgnt2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgnt2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgnt2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgnt2\_6/prodata/2/1aa/PCUTS\_COMB.pep:\*  
6: /cgnt2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3927.5	77.3	925	1	US-08-252-995D-4
2	3927.5	77.3	925	2	US-08-834-108-4
3	1926	37.9	464	1	US-08-252-995D-6
4	1926	37.9	464	2	US-08-834-108-6
5	1883.5	37.1	416	1	US-08-252-995D-2
6	1883.5	37.1	416	2	US-08-834-108-2
7	1370	27.0	273	1	US-08-252-995D-10
8	1370	27.0	273	2	US-08-834-108-10
9	590.5	11.6	607	2	US-08-878-989-15
10	590.5	11.6	607	3	US-09-272-796-15
11	578.5	11.4	271	1	US-08-252-995D-11
12	578.5	11.4	271	2	US-08-834-108-11
13	561.5	11.1	272	1	US-08-252-995D-12
14	561.5	11.1	272	2	US-08-834-108-12
15	560.5	11.0	685	2	US-08-878-989-1
16	560.5	11.0	685	3	US-09-136-282-2
17	560.5	11.0	685	3	US-09-272-796-1
18	560.5	11.0	685	3	US-09-505-744-2
19	545	10.7	603	4	US-09-311-311C-26
20	533	10.5	603	3	US-09-198-122-2
21	525.5	10.3	272	1	US-08-252-995D-14
22	525.5	10.3	272	2	US-08-834-108-14
23	508.5	10.0	403	2	US-08-755-728-4
24	508.5	10.0	403	3	US-08-974-655-4
25	508.5	10.0	403	3	US-09-283-011-4
26	498	9.8	722	4	US-08-817-8328-32
27	495.5	9.8	722	4	US-09-984-890-4

28	490	9.6	275	1	US-08-252-995D-13	Sequence 13, App1
29	490	9.6	275	2	US-08-834-108-13	Sequence 13, App1
30	487	9.6	724	4	US-09-984-890-2	Sequence 2, App1
31	485.5	9.6	344	2	US-08-755-728-3	Sequence 3, App1
32	485.5	9.6	344	2	US-08-974-655-3	Sequence 3, App1
33	485.5	9.6	344	3	US-09-283-011-3	Sequence 3, App1
34	480	9.5	347	2	US-09-016-000-1	Sequence 1, App1
35	479.5	9.4	745	4	US-09-523-849-36	Sequence 36, App1
36	464	9.1	729	2	US-08-677-298-2	Sequence 2, App1
37	464	9.1	729	4	US-09-523-849-33	Sequence 33, App1
38	462.5	9.1	633	3	US-08-557-006C-43	Sequence 43, App1
39	454.5	9.0	556	2	US-09-016-000-4	Sequence 4, App1
40	454.5	9.0	556	4	US-09-156-793D-2	Sequence 2, App1
41	450	8.9	1037	4	US-09-428-711A-21	Sequence 21, App1
42	449	8.8	793	4	US-09-523-849-32	Sequence 32, App1
43	446.5	8.8	556	4	US-09-800-960-4	Sequence 4, App1
44	446.5	8.8	556	4	US-10-096-960-4	Sequence 4, App1
45	445.5	8.8	1203	4	US-09-799-875-5	Sequence 5, App1

#### ALIGNMENTS

RESULT 1  
US-08-252-995D-4  
Sequence 4, Application US/08252995D  
Patent No. 650501  
GENERAL INFORMATION:  
APPLICANT: Demille, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3J2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdyak, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 925 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-995D-4  
Query Match 77.3%; Score 3927.5; DB 1; Length 925;  
Best Local Similarity 78.6%; Pred. No. 8.4e-296;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;  
QY 1 MATCGEKEDPKVGNLJGKSPAGYRAESHTGIEVAIKIMDKAMTKAGVQVQV 60  
DB 1 MACICGIEDPKVGNLJGKSPAGYRAESHTGIEVAIKIMDKAMTKAGVQVQV 60  
QY 61 VKIHGQLKPSIETLVYVFEQSGEMRNYLKNRVKVPFSEARHFMHQI 120

```

Db      61 VKHCOLKHPSEVLKLYNFEDNNVYLVLEKCHNGENRKYKMKCFSEBEAHFMHQI 120
Qy      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Db      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Qy      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVTKNTLKVVLADYEMPSFLIRAKD 240
Db      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVTKNTLKVVLADYEMPSFLIRAKD 240
Qy      241 LIHOLLRNPNADRLSLSSVLDHPMSRNSSTKSDLTGTVEDSIDSGHATTSTAITASST 300
Db      241 LIHOLLRNPNADRLSLSSVLDHPMSRNSSTKSDLTGTVEDSIDSGHATTSTAITASST 300
Qy      301 SLSGSLD-RRLVAGQULPKKITVFOKNKNSDP--SSGDSNFCOTOMGNPEQENRGRG 358
Db      301 SLSGSLD-RRLVAGQULPKKITVFOKNKNSDP--SSGDSNFCOTOMGNPEQENRGRG 358
Qy      359 RVIODAERPHSRILRAAYSDDSGTNSQSQAATYMERCHSAEMLSVKRSQGGGENEE 418
Db      359 RVIODAERPHSRILRAAYSDDSGTNSQSQAATYMERCHSAEMLSVKRSQGGGENEE 418
Qy      419 RYSPDTNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGKTPPPFPADPTOTETVQO 478
Db      419 RYSPDTNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGKTPPPFPADPTOTETVQO 478
Qy      479 WFGNLOINAHLRKTEYDSISPNRDFQGHDPLOKOTSKNAMTDTKYKNSDASDNASHYK 538
Db      479 WFGNLOINAHLRKTEYDSISPNRDFQGHDPLOKOTSKNAMTDTKYKNSDASDNASHYK 538
Qy      539 QQNTMTKMTALHSKPELIIQOECVFGSDPLSEOSKTRGMEPPMGYQNRRLRSITSPVAHR 598
Db      539 QQNTMTKMTALHSKPELIIQOECVFGSDPLSEOSKTRGMEPPMGYQNRRLRSITSPVAHR 598
Qy      599 LKRIROTKKAVYSIIDSEVCYELVEXASOEYVEVLOISSDGMTITTYNNGRGRRP 658
Db      599 LKRIROTKKAVYSIIDSEVCYELVEXASOEYVEVLOISSDGMTITTYNNGRGRRP 658
Qy      659 LADRPSPFDNISRYSPFDMLEKYMRYOYASRFVOLVRSKPKITTYFTRAKCLIMENS 718
Db      659 LADRPSPFDNISRYSPFDMLEKYMRYOYASRFVOLVRSKPKITTYFTRAKCLIMENS 718
Qy      719 PGADFEVWYDVYKIKHTDFOVIEKTKSYTLKSESEVNIKEETIKMYMDHANGRI 778
Db      719 PGADFEVWYDVYKIKHTDFOVIEKTKSYTLKSESEVNIKEETIKMYMDHANGRI 778
Qy      779 CLALBSIIEBERKTSAPFPFIIIGRKGSTSPKALSPPSVDSNYPTRDRASPNRV 838
Db      779 CLALBSIIEBERKTSAPFPFIIIGRKGSTSPKALSPPSVDSNYPTRDRASPNRV 838
Qy      839 MHSASPTQAPILINPMVTNEGILTTASGTDISSNLKDCLPKSAQLLKSVFVYVGM 898
Db      839 MHSASPTQAPILINPMVTNEGILTTASGTDISSNLKDCLPKSAQLLKSVFVYVGM 898
Qy      899 ATQOLTSAGVWVQNDOSQLVQAGVSSISYTSPNGOTPRYGENEKLPDVITKLOCLSSI 958
Db      899 ATQOLTSAGVWVQNDOSQLVQAGVSSISYTSPNGOTPRYGENEKLPDVITKLOCLSSI 958
Qy      959 LLMFSNPTNPF 969
Db      914 LLMFSNPTNPF 924

```

RESULT 2  
US-08-834-108-4  
; Sequence 4, Application US/08834108  
; Patent No. 5976893

GENERAL INFORMATION:  
; APPLICANT: Dennis, James W  
; APPLICANT: Heffernan, Mike  
; APPLICANT: Fode, Carol  
; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE

```

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,108
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydk, Linda M
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-834-108-4

```

```

Query Match      77.3%; Score 3927.5; DB 2; Length 925;
Best Local Similarity 78.6%; Pred. No. 8.4e-296;
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;

Qy      1 NATIGEKIEDPFKVNLLGKSPAGVYRAESIHGLEVAIMIDKKAMTKAMTVQVONE 60
Db      1 MACIGERIEDPFKVNLLGKSPAGVYRAESIHGLEVAIMIDKKAMTKAMTVQVONE 60
Qy      61 VKHCOLKHPSEVLKLYNFEDNNVYLVLEKCHNGENRKYKMKCFSEBEAHFMHQI 120
Db      61 VKHCOLKHPSEVLKLYNFEDNNVYLVLEKCHNGENRKYKMKCFSEBEAHFMHQI 120
Qy      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Db      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Qy      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVTKNTLKVVLADYEMPSFLIRAKD 240
Db      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVTKNTLKVVLADYEMPSFLIRAKD 240
Qy      241 LIHOLLRNPNADRLSLSSVLDHPMSRNSSTKSDLTGTVEDSIDSGHATTSTAITASST 300
Db      241 LIHOLLRNPNADRLSLSSVLDHPMSRNSSTKSDLTGTVEDSIDSGHATTSTAITASST 300
Qy      301 SLSGSLD-RRLVAGQULPKKITVFOKNKNSDP--SSGDSNFCOTOMGNPEQENRGRG 358
Db      301 SLSGSLD-RRLVAGQULPKKITVFOKNKNSDP--SSGDSNFCOTOMGNPEQENRGRG 358
Qy      359 RVIODAERPHSRILRAAYSDDSGTNSQSQAATYMERCHSAEMLSVKRSQGGGENEE 418
Db      359 RVIODAERPHSRILRAAYSDDSGTNSQSQAATYMERCHSAEMLSVKRSQGGGENEE 418
Qy      419 RYSPDTNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGKTPPPFPADPTOTETVQO 478
Db      419 RYSPDTNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGKTPPPFPADPTOTETVQO 478
Qy      479 WFGNLOINAHLRKTEYDSISPNRDFQGHDPLOKOTSKNAMTDTKYKNSDASDNASHYK 538
Db      479 WFGNLOINAHLRKTEYDSISPNRDFQGHDPLOKOTSKNAMTDTKYKNSDASDNASHYK 538
Qy      539 QQNTMTKMTALHSKPELIIQOECVFGSDPLSEOSKTRGMEPPMGYQNRRLRSITSPVAHR 598

```

Db	502	QUSAKYVSAAHHKKEVWFOEP - GLHPHSBOSKNSMESTLQYQKPTLRSTISPLAHR	5599
Qy	599	LKPIRQTKKAVSILDSBEVCEVELKEVASOBVYKEVLIQISSDQNTITITYENGRGFP	6588
Db	560	LKPIRQTKKAVSILDSBEVCEVELLRKCASEGIVKEVLIQISSDQNTITITYENDDRGFP	619
Qy	659	LADRPSPSTDNISRYSPDNLPEKYWRKQYASRFVQVLVRSKSPKITYPTRYAKCIIMENS	718
Db	620	LADRPPLPTDINISRYSPDNLPEKYWRKQYASRFQVLVRSKPKITYPTTRYACIIMENS	679
Qy	719	PGADPEWVFYQGVCKHKTEDFIQVLEKQKGSYTLKSESEVNSLKEEIKYMDHANGHR	778
Db	680	PGADPEWVFYQGVCAKHKTEENLHILEKGI SYNLKNEEVTSLKEEIKYMDHANGHR	739
Qy	779	CLALESIISSEEBKRSAFPFIIIGRKGGSTSPALSPPVSDNSYENPTRDASFNRMV	838
Db	740	CLLSLESVISSEEBKRSGSFPIIYGRKKGNTSPFALNAPF - VDSCKCGEBAASRLS	798
Qy	839	MHSAASPQAPILNDSMYTNEGILGTTTASGTDISGNSLKDCLPKSAQILLKSVFVNVGV	898
Db	799	VNSAFAFPQSPCLSPSTVYVEGLGHATATATGVSSS-----LPKSAQILLKSVFVNVGV	853
Qy	899	ATQLTSGAVWQFNDGSQLVWQAGVSSISYTSBNGCTTYGGENEKLPTYIKQQLCLSSI	958
Db	854	ATQLTSGAVWQFNDGSQLVWQAGVSSISYTSBQCTTYGGENEKLPEYIKQQLCLSSI	913
Qy	959	LLMFNSPTPNF	969
Db	914	LLMFNSPTPNF	924

```

1      RESULT 3 252-995D-6
2      US-08-252-995D-6
3      Sequence 6, Application US/08252995D
4      Patent No. 5650501
5      GENERAL INFORMATION:
6      APPLICANT: Dennie, James W
7      APPLICANT: Heffernan, Mike
8      APPLICANT: Fode, Carol
9      TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
10     NUMBER OF SEQUENCES: 14
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: BERSKIN & PARR
13     STREET: 40 King Street West
14     CITY: Toronto
15     STATE: Ontario
16     COUNTRY: Canada
17     ZIP: M5H 3Y2
18
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: Patentin Release #1.0, Version #1.30
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/252,995D
26     FILING DATE: 02-JUN-1994
27     CLASSIFICATION: 536
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Kurdavsk, Linda M
30     REGISTRATION NUMBER: 34,971
31     REFERENCE/DOCKET NUMBER: 3153-96
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: (416) 364-7311
34     TELEFAX: (416) 361-1398
35     INFORMATION FOR SEQ. ID NO. 6:
36     SEQUENCE CHARACTERISTICS:
37     LENGTH: 464 amino acids
38     TYPE: amino acid
39     TOPOLOGY: linear
40     MOLECULE TYPE: protein
41     US-08-252-995D-6

```

Query Match	37.9%	Score 1926;	DB 1;	Length 464;
Best Local Similarity	84.9%	Pred. No. 5.1e-141;		
Matches 372;	Conservative 30;	Mismatches 30;	Indels 6;	Gaps 5;

Qy	1	MATIGEXIEBPKGNLLGKGSFAGVPAEISHTGL EVAIKMIDKKAMKAMQVRQNE	60
Dp	1	MAATIGERIEBPKGNLLGKGSFAGVPAEISHTGL EVAIKMIDKKAMKAMQVRQNE	60
Qy	61	VKTHCOLKHPISLELYNFYEDSNYYVYLEMCHENGEMNRYYLKNVYKPSSENAHPHMOI	120
Dp	61	VKTHCOLKHPISLELYNFYEDSNYYVYLEMCHENGEMNRYYLKNRNMKPFSEARARHPHMOI	120
Qy	121	ITGMVLYHSHSHTLRDFTLSNILLTRNNNKIKADGGLATOLKMPHEKHTLGGTPYISP	180
Dp	121	ITGMVLYHSHSHTLRDFTLSNILLTRNNNKIKADGGLITOLMMPHEKHTLGGTPYISP	180
Qy	181	EIATRSAGLESDDVWSLQCMFYTLIGRPPTDITVKNITLNVVLADYEMPSFLSTBAQD	240
Dp	181	EIATRSAGLESDDIWSJGCMGYTLIGRPPTDITVKNITLNVVLADYEMPAFLSREAO	240
Qy	241	LIIHOLLRRNPDRSLSSVLDHPFMSRNRSSITKXKDLGTVEEDSIDSGHATISPAITASSST	300
Dp	241	LIIHOLLRRNPDRSLSSVLDHPFMSRNPSPKSKXGVTEEDSMDSGHATISPAITASSST	300
Qy	301	SISGSLFPIKRLTLIGQPLFNQWYVYPPKKKSGSTDFSSSDGNSPFTQWGN--OETSNSGSG	358
Dp	301	SISGSLIAD--RLLVQGPLPKITTFQKKKNSDF--SSDGSNFCYQWGNPBEAANSRGSG	358
Qy	359	RVIODAEBRPHSRYYLRRAYSSDSRGTSNSQSQAKTYTWERCHSAEMLSVSKSGGGENEE	418
Dp	359	RVIDEAEBRPHSRYYLRRAYSSDDRAPSPN--QSRAKTYSEVERCHSIVEMLSKTPRRS--LDENOH	416
Qy	419	RYSPTDNNANIFNFPEK 436	
Dp	417	RYSPTKSNVNVLTSLNTK 434	

RESULT 4  
 US-08-834-108-6  
 Sequence 6, Application US/08834108  
 Patent No. 5976893  
 GENERAL INFORMATION:  
 APPLICANT: Dennis, James W  
 APPLICANT: Heffernan, Mike  
 APPLICANT: Fode, Carol  
 TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BERSKIN & PARR  
 STREET: 40 King Street West  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/834,108  
 FILING DATE:  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kurydyk, Linda M  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-210  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 464 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-834-108-6

Query Match 37.9%; Score 1926; DB 2; Length 464;  
Best Local Similarity 84.9%; Pred. No. 5.1e-141;  
Matches 372; Conservative 30; Mismatches 30; Indels 6; Gaps 5;

QY 1 MATCIGKIDPKVGNLKGSPAGVYRAESIHTELEVALIKMIDKKAMYKAGVQVONE 60  
DB 1 MAACIGRIEDPKVGNLKGSPAGVYRAESIHTELEVALIKMIDKKAMYKAGVQVONE 60  
QY 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRILKNRKVPSEARHFMHOI 120  
DB 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRILKNRKVPSEARHFMHOI 120  
QY 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLGTPNYISP 180  
DB 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLGTPNYISP 180  
QY 181 EATRSAGLESVDWSLGCMPYTLILGRPPDTDTVKNLTKVLADEMPFLSTEAKD 240  
DB 181 EATRSAGLESVDWSLGCMPYTLILGRPPDTDTVKNLTKVLADEMPFLSTEAKD 240  
QY 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATITASSST 300  
DB 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATITASSST 300  
QY 301 SIISGLFDRKRLILGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
DB 301 SIISGLD-RRLLVGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
QY 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSOQATYTMERCHSAEMLSVSKSGGGENE 418  
DB 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSOQATYTMERCHSAEMLSVSKSGGGENE 418  
QY 419 RYSPPTNNANIFPFKEK 436  
DB 419 RYSPPTNNANIFPFKEK 436  
QY 417 RYSPPTNNANIFPFKEK 434  
DB 417 RYSPPTNNANIFPFKEK 434

## RESULT 5

US-08-252-995D-2  
Sequence 2, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
APPLICANT: Fode, Carol  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 416 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-995D-2

Query Match 37.1%; Score 1883.5; DB 1; Length 416;  
Best Local Similarity 87.7%; Pred. No. 8.6e-138;  
Matches 362; Conservative 27; Mismatches 19; Indels 5; Gaps 4;

QY 1 MATCIGKIDPKVGNLKGSPAGVYRAESIHTELEVALIKMIDKKAMYKAGVQVONE 60  
DB 1 MAACIGRIEDPKVGNLKGSPAGVYRAESIHTELEVALIKMIDKKAMYKAGVQVONE 60  
QY 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRILKNRKVPSEARHFMHOI 120  
DB 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRILKNRKVPSEARHFMHOI 120  
QY 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLGTPNYISP 180  
DB 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLGTPNYISP 180  
QY 181 EATRSAGLESVDWSLGCMPYTLILGRPPDTDTVKNLTKVLADEMPFLSTEAKD 240  
DB 181 EATRSAGLESVDWSLGCMPYTLILGRPPDTDTVKNLTKVLADEMPFLSTEAKD 240  
QY 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATITASSST 300  
DB 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATITASSST 300  
QY 301 SIISGLFDRKRLILGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
DB 301 SIISGLD-RRLLVGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
QY 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSOQATYTMERCHSAEMLSVSKSGGGENE 411  
DB 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSOQATYTMERCHSAEMLSVSKSGGGENE 411

## RESULT 6

US-08-834-108-2  
Sequence 2, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
APPLICANT: Fode, Carol  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 416 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-834-108-2

Query Match 37.1%; Score 1883.5; DB 2; Length 416;  
Best Local Similarity 87.7%; Pred. No. 8.6e-138;  
Matches 362; Conservative 27; Mismatches 19; Indels 5; Gaps 4;

QY 1 MATCGEKEDPKVGNLGGSPAGYRAESHTGAEVAIKMIDKAMKAGVORVONE 60  
DB 1 MACIGERIEDPKVGNLGGSPAGYRAESHTGAEVAIKMIDKAMKAGVORVONE 60  
QY 61 VTIHQKHPSTILELYNYPEDSNVYVLVLEMCNGEMNRYLKORVVPSENEARHFMHQI 120  
DB 61 VTIHQKHPSTILELYNYPEDSNVYVLVLEMCNGEMNRYLKORVVPSENEARHFMHQI 120  
QY 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATQOLKMPHEKHVTLGCTPNYISP 180  
DB 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATQOLNMPHEKHVTLGCTPNYISP 180  
QY 181 EIATSSAGLESVDWMLGCMFTYLLIGRPEDTDYVNTLNKVVADYEMPSFLSEAKD 240  
DB 181 EIATSSAGLESVDWMLGCMFTYLLIGRPEDTDYVNTLNKVVADYEMPSFLSEAKD 240  
QY 241 LHHLLRRPADRLSLSSVLDHPFMSRNSSTKSDIGTEDSIDSGHATITSTITASST 300  
DB 241 LHHLLRRPADRLSLSSVLDHPFMSRNSSTKSDIGTEDSIDSGHATITSTITASST 300  
QY 301 SISGLDFDKRRLILGQPLPNKMTVPKNSSTDPSSSGDGNSFTYTMGN--QETNSGRG 358  
DB 301 SISGLDFDKRRLILGQPLPNKMTVPKNSSTDPSSSGDGNSFTYTMGN--QETNSGRG 358  
QY 359 RVIQDAEERPHSRYLRRAYSSDRSGTSNQSQAQYTTMERCHEAELYSKRS 411  
DB 359 RVIQDAEERPHSRYLRRAYSSDRSGTSNQSQAQYTTMERCHEAELYSKRS 411

RESULT 7  
US-08-252-995D-10  
Sequence 10, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-252-995D-10

Query Match 27.0%; Score 1370; DB 1; Length 273;  
Best Local Similarity 94.5%; Pred. No. 3e-98;  
Matches 258; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 5 IGEKIEDPKVGNLGGSPAGYRAESHTGAEVAIKMIDKAMKAGVORVONEVKIH 64  
DB 1 IGERIEDPKVGNLGGSPAGYRAESHTGAEVAIKMIDKAMKAGVORVONEVKIH 60  
QY 65 COLKHPSTILELYNYPEDSNVYVLVLEMCNGEMNRYLKORVVPSENEARHFMHQIITGM 124  
DB 65 COLKHPSTILELYNYPEDSNVYVLVLEMCNGEMNRYLKORVVPSENEARHFMHQIITGM 120  
QY 125 LTIHSHGILHRDLTSLNLLTRNNMIKIADFGIATQOLKMPHEKHVTLGCTPNYISP 184  
DB 125 LTIHSHGILHRDLTSLNLLTRNNMIKIADFGIATQOLNMPHEKHVTLGCTPNYISP 180  
QY 185 RSHAGLESVDWMLGCMFTYLLIGRPEDTDYVNTLNKVVADYEMPSFLSEAKD 244  
DB 185 RSHAGLESVDWMLGCMFTYLLIGRPEDTDYVNTLNKVVADYEMPSFLSEAKD 240  
QY 245 LRRNPADRLSLSSVLDHPFMSRNSSTKSDIG 277  
DB 245 LRRNPADRLSLSSVLDHPFMSRNSSTKSDIG 273

RESULT 8  
US-08-834-108-10  
Sequence 10, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398



```
/ APPLICATION NUMBER: US/09/272,796
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/878,989
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J J
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0321 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 607 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 1827450
/ US-09-272-796-15

Query Match 11.6%; Score 590.5; DB 3; Length 607;
Best Local Similarity 37.5%; Pred. No. 3,1e-37;
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNILGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 74
DB GLLLGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 74
QY 26 GLLLGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 85
DB GLLLGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 85
QY 75 LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 134
DB LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 134
QY 86 FSHHEDADNIIYFELCKRSKSLAHMKAR-HTLLEPEVRYRQLSGKYLHOGILH 144
DB FSHHEDADNIIYFELCKRSKSLAHMKAR-HTLLEPEVRYRQLSGKYLHOGILH 144
QY 135 RDLTSLNLLTTRNMNIIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGLESDV 194
DB RDLTSLNLLTTRNMNIIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGLESDV 194
QY 145 RLKLGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 204
DB RLKLGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 204
QY 195 WSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 254
DB WSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 254
QY 205 WSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 264
DB WSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 264
QY 255 SLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 304
DB SLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 304
QY 265 SLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 310
DB SLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 310
QY 305 SLFDRKR 311
DB SLFDRKR 311
QY 311 SLFGRKK 317
DB SLFGRKK 317

RESULT 11
US-08-252-995D-11
/ Sequence 11, Application US/08252995D
/ Patent No. 5650501
/ GENERAL INFORMATION:
/ APPLICANT: Dennis, James W
/ APPLICANT: Heffernan, Mike
/ APPLICANT: Fode, Carol
/ TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BERSKIN & PARR
/ STREET: 40 King Street West
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5H 3Y2
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
```

```
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/252,995D
/ FILING DATE: 02-JUN-1994
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kurydyk, Linda M
/ REGISTRATION NUMBER: 34,971
/ REFERENCE/DOCKET NUMBER: 3153-96
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 364-7311
/ TELEFAX: (416) 361-1398
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 271 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Drosophila melanogaster
/ US-08-252-995D-11

Query Match 11.4%; Score 578.5; DB 1; Length 271;
Best Local Similarity 42.5%; Pred. No. 7,4e-37;
Matches 108; Conservative 53; Mismatches 92; Indels 1; Gaps 1;

QY 12 FVGNILGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 71
DB FVGNILGKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 71
QY 7 YKMRPFKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 66
DB YKMRPFKSPAGYRAESHTGLEVALIKMIDKAMYGAVQVQVONEVKIHCQKPSILE 66
QY 72 LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 131
DB LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 131
QY 67 LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 125
DB LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 125
QY 133 RDLTSLNLLTTRNMNIIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGLESDV 191
DB RDLTSLNLLTTRNMNIIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGLESDV 191
QY 126 LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 185
DB LNVYFEDSNVYVLEMCNGENRRLKRVKVPFSENEARHFMHOITGMLYHSHG 185
QY 192 SDVSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 251
DB SDVSLGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 251
QY 186 VDWISGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 245
DB VDWISGCMFYTLIGRPEDTDVYKNTLVKLVADYEMPSFLSIKADLIHQLLRNPA 245
QY 252 DRLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 304
DB DRLSSVLDHFPFMRNSSTKSDGTVDSDSGHATITSLTASSST-----SIG 304
QY 246 SRPAIGQLNFEFL 259
DB SRPAIGQLNFEFL 259

RESULT 12
US-08-834-108-11
/ Sequence 11, Application US/08834108
/ Patent No. 5976893
/ GENERAL INFORMATION:
/ APPLICANT: Dennis, James W
/ APPLICANT: Heffernan, Mike
/ APPLICANT: Fode, Carol
/ TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BERSKIN & PARR
/ STREET: 40 King Street West
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5H 3Y2
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Drosophila melanogaster  
US-08-834-108-11

Query Match 11.4%; Score 578.5; DB 2; Length 271;  
Best Local Similarity 42.5%; Pred. No. 7,4e-37;  
Matches 108; Conservative 53; Mismatches 92; Indels 1; Gaps 1;

QY 12 PKVGNLGGKSPAGVYRAESHTGLEVAIKMTDKKAMVQVQVONEVYKIHCOLKPS 71  
DB 7 YKRMRFPGGGRKACEITIDVEDVDPAKIVSKMLMHNQEKRAQETTHRSLNHN 66  
QY 72 ILELVNPFDSNVYVLEMCNCHGEMNRYLKNRVKPSSENAHFMHQIITGMVLYSHG 131  
DB 67 IYVFNHYFSDSQIYIVLELCKRSMMELHRR-KSITEFCRYIYQIIGVXYLHNR 125  
QY 132 IAHRLTSLNLTTRMNTKIADFGLATQKMPHEKHTLCTPNVISPETIARSAGLE 191  
DB 126 ITHRDKLNLLFNDLHVKIGPGLATRIEYGEKKTLCGTANYIAPILTYKGGHSE 185  
QY 192 SDVWSLGCNFFYLLIGRPFDPTVYKNTLKVVLADYEMPSFLSEAKDLIHLRRNPA 251  
DB 186 VDIWSIGCMYTLVQPFETKTLADTYSKIKCKEYRPSYLRKRAADNVIMLQPNRE 245  
QY 252 DRLSSVLDHPM 265  
DB 246 SRPAIGQLNFEFL 259

RESULT 13  
US-08-252-995D-12  
Sequence 12, Application US/08252995D  
Patent No. 5650501

GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARK  
STREET: 40 King Street West  
STATE: Ontario  
CITY: Toronto  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:

NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-252-995D-12

Query Match 11.1%; Score 561.5; DB 1; Length 272;  
Best Local Similarity 41.6%; Pred. No. 1,6e-35;  
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNULGGKSPAGVYRAESHTGLEVAIKMTDKKAMVQVQVONEVYKIHCOLKPS 74  
DB 11 GAVLGGKSPAGVYRAESHTGLEVAIKMTDKKAMVQVQVONEVYKIHCOLKPS 70  
QY 75 LVNPFDSNVYVLEMCNCHGEMNRYLKNRVKPSSENAHFMHQIITGMVLYSHG 134  
DB 71 FYHVPEDKNYIYLLFECGRSMHILKAR-KVLEPEVRYVLRQVSGKTLHEGEI 129  
QY 135 RDLTSLNLTTRMNTKIADFGLATQKMPHEKHTLCTPNVISPETIARSAGLE 194  
DB 130 RDLKGNFPIINAMELVKDFGLAARLEPLERRRTICTPNVLSPEVLNKGHGESDI 189  
QY 195 WSLGCMFFYLLIGRPFDPTVYKNTLKVVLADYEMPSFLSEAKDLIHLRRNPA 254  
DB 190 WADGCMYTLVQPFETKTLADTYSKIKCKEYRPSYLRKRAADNVIMLQPNRE 249  
QY 255 SLSSVLDHPF 264  
DB 250 SLDDIRHDF 259

RESULT 14  
US-08-834-108-12  
Sequence 12, Application US/08834108  
Patent No. 5976893

GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARK  
STREET: 40 King Street West  
STATE: Ontario  
CITY: Toronto  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: September 24, 2004, 14:49:50 ; Search time 198 Seconds  
(without alignments)  
8164.505 Million cell updates/sec

Title: US-10-026-021-1

Sequence: 1 atggcgacctgcgcgcgggga.....cgaccttaatttcattga 2913

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3937	4 US-09-620-312D-280	Sequence 280, App1
2	1879	64.5	3447	1 US-08-252-995D-3	Sequence 3, App1
3	1879	64.5	3447	2 US-08-834-108-3	Sequence 3, App1
4	970.4	33.3	1600	1 US-08-252-995D-5	Sequence 5, App1
5	970.4	33.3	1600	2 US-08-834-108-5	Sequence 5, App1
6	924.2	31.7	1453	1 US-08-252-995D-1	Sequence 1, App1
7	924.2	31.7	1453	2 US-08-834-108-1	Sequence 1, App1
8	154.2	5.3	381	4 US-09-016-434-649	Sequence 649, App1
9	134.4	4.6	2770	2 US-08-878-989-8	Sequence 8, App1
10	134.4	4.6	2770	3 US-09-272-796-8	Sequence 8, App1
11	134.4	4.6	2789	3 US-09-136-282-3	Sequence 3, App1
12	134.4	4.6	2789	3 US-09-505-744-3	Sequence 3, App1
13	132.4	4.5	2783	3 US-09-136-282-1	Sequence 1, App1
14	132.4	4.5	2783	3 US-09-505-744-1	Sequence 1, App1
15	131.4	4.5	2169	4 US-09-016-434-1147	Sequence 1147, App1
16	118	4.1	2033	4 US-09-660-925B-10	Sequence 10, App1
17	113.2	3.9	2198	2 US-08-755-728-2	Sequence 2, App1
18	113.2	3.9	2198	2 US-08-974-655-2	Sequence 2, App1
19	113.2	3.9	2198	3 US-09-283-011-2	Sequence 2, App1
20	113.2	3.9	2253	4 US-09-660-925B-3	Sequence 3, App1
21	109.4	3.8	1244	2 US-09-198-122-1	Sequence 1, App1
22	104	3.6	1244	2 US-08-755-728-1	Sequence 1, App1
23	104	3.6	1244	2 US-08-974-655-1	Sequence 1, App1
24	104	3.6	1244	2 US-09-283-011-1	Sequence 1, App1
25	103.6	3.6	1244	2 US-09-359-161-1	Sequence 1, App1
26	99	3.4	1891	3 US-09-489-466-1	Sequence 1, App1
27	97.4	3.3	1864	4 US-09-156-793D-1	Sequence 1, App1

28	97.4	3.3	1890	3 US-09-289-466-2	Sequence 2, App1
29	97.4	3.3	1929	2 US-09-016-000-10	Sequence 10, App1
30	96	3.3	1281	2 US-09-016-000-7	Sequence 7, App1
31	89.8	3.1	1590	4 US-09-554-726A-4	Sequence 4, App1
32	89.8	3.1	1732	4 US-09-430-564-1	Sequence 1, App1
33	89.8	3.1	1257	4 US-09-554-726A-6	Sequence 6, App1
34	89.2	3.1	2556	4 US-09-817-310-1	Sequence 1, App1
35	87.8	3.0	2754	3 US-09-429-322-3	Sequence 3, App1
36	87	3.0	1637	2 US-08-966-316-10	Sequence 10, App1
37	86	3.0	4438	4 US-09-566-921-81	Sequence 81, App1
38	82.4	2.8	2698	2 US-08-677-298-1	Sequence 1, App1
39	82.2	2.8	2076	4 US-09-554-726A-11	Sequence 11, App1
40	82.2	2.8	2827	4 US-09-554-726A-11	Sequence 11, App1
41	82.2	2.8	2827	4 US-09-554-726A-20	Sequence 20, App1
42	80.6	2.8	2298	4 US-09-975-326-3	Sequence 3, App1
43	80.6	2.8	2298	4 US-10-217-357-3	Sequence 3, App1
44	80.6	2.8	2301	4 US-09-975-326-1	Sequence 1, App1
45	80.6	2.8	2301	4 US-10-217-357-1	Sequence 1, App1

## ALIGNMENTS

RESULT 1  
US-09-620-312D-280  
; Sequence 280, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyang  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonhong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungang  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinheast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 656962e1 Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 280  
; LENGTH: 3937  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (334)..(3246)  
; US-09-620-312D-280

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGCGGAGAAATCGAGATTGTAAGTTGGAATCTGCTGTGTA 60  
DB 334 ATGGCGACCTGCATCGCGGAGAAATCGAGATTGTAAGTTGGAATCTGCTGTGTA 393

QY 61 GAATCATTTGCTGTGTCTACAGAGCTAGTCCATTGACATCTGGTTTGAAGTTGCAATC 120  
DB 394 GGAATCATTTGCTGTGTCTACAGAGCTAGTCCATTGACATCTGGTTTGAAGTTGCAATC 453  
QY 121 AAAATGATTAATAAGAAAGCCATGTTACAAAGCAGAAATGCTTACAGAGTCCAAAATGAG 180  
DB 454 AAAATGATTAATAAGAAAGCCATGTTACAAAGCAGAAATGCTTACAGAGTCCAAAATGAG 513  
QY 181 GTGAAAAATACATTTGSCCAATTTGAAACATCTTTCTATCTGTGGAGCTTTATTAATACTATTTTGA 240  
DB 514 GTGAAAAATACATTTGSCCAATTTGAAACATCTTTCTATCTGTGGAGCTTTATTAATACTATTTTGA 573  
QY 241 GATAGCAATTAATGTATCTGTATTAAGAAATGTCCATTAATGAGAAATGAACAGGTAT 300  
DB 574 GATAGCAATTAATGTATCTGTATTAAGAAATGTCCATTAATGAGAAATGAACAGGTAT 633  
QY 301 CTAAAGAAATGAGTGAACCCCTTCTGAGAAATGAAAGCTGCACTTTGATGCAACAGATC 360  
DB 634 CTAAAGAAATGAGTGAACCCCTTCTGAGAAATGAAAGCTGCACTTTGATGCAACAGATC 693  
QY 361 ATGACAGGAGTGTGTATCTTCAATCTCATGTATTAATGACACCGGAGCTCAACATTTCT 420  
DB 694 ATGACAGGAGTGTGTATCTTCAATCTCATGTATTAATGACACCGGAGCTCAACATTTCT 753  
QY 421 AACCTCTTACTGACTGTAATATGAACATCAAGATTCGATTTTGGGCTGGCACTCAA 480  
DB 754 AACCTCTTACTGACTGTAATATGAACATCAAGATTCGATTTTGGGCTGGCACTCAA 813  
QY 481 CTGAAAAATGCAATGAAAGCACTATACATTAATGTGGAATCTCTTAATCAATTTTCAACA 540  
DB 814 CTGAAAAATGCAATGAAAGCACTATACATTAATGTGGAATCTCTTAATCAATTTTCAACA 873  
QY 541 GAAATTTGCACTGGAAGTGCATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600  
DB 874 GAAATTTGCACTGGAAGTGCATGAGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 933  
QY 601 TTTTATATCAATTAATATGCGGAGAACACCTTTGCACTGCAACAGTCAAGAACATTTA 660  
DB 934 TTTTATATCAATTAATATGCGGAGAACACCTTTGCACTGCAACAGTCAAGAACATTTA 993  
QY 661 AATAAAGTATATGSCAGATTAATGAATGCAATCTTTTGTGCAATTAAGGCCAAGAAC 720  
DB 994 AATAAAGTATATGSCAGATTAATGAATGCAATCTTTTGTGCAATTAAGGCCAAGAAC 1053  
QY 721 CTATATTCACAGTTACTTGTGAGAAATTCAGCAGATCGTTTAACTGTCTTCAATATG 780  
DB 1054 CTATATTCACAGTTACTTGTGAGAAATTCAGCAGATCGTTTAACTGTCTTCAATATG 1113  
QY 781 GACCAATCCTTTATATGCCCCGAAATCTTCAACAAAAAGTAAAGATTTAGAACTGTGAA 840  
DB 1114 GACCAATCCTTTATATGCCCCGAAATCTTCAACAAAAAGTAAAGATTTAGAACTGTGAA 1173  
QY 841 GACTCAATTAATGAGGAGCATGCCAATTTCTACTGCAATTAACAGCTTCTTCAAGTACC 900  
DB 1174 GACTCAATTAATGAGGAGCATGCCAATTTCTACTGCAATTAACAGCTTCTTCAAGTACC 1233  
QY 901 AGTATTAAGTGTATGTTTATTTGAACAAAGAGACTTTGATGTCAGCCACTCCCAAT 960  
DB 1234 AGTATTAAGTGTATGTTTATTTGAACAAAGAGACTTTGATGTCAGCCACTCCCAAT 1293  
QY 961 AAAATGACTGTATTTCCAAAGAAATAAAGTCAATGATTTTCTTCTTCAAGAGATGGA 1020  
DB 1294 AAAATGACTGTATTTCCAAAGAAATAAAGTCAATGATTTTCTTCTTCAAGAGATGGA 1353  
QY 1021 AACAGTTTATTAATCAAGTGGGAAATCAAGAAACCAAGTAATAGTGAAGGGAGAGATA 1080  
DB 1354 AACAGTTTATTAATCAAGTGGGAAATCAAGAAACCAAGTAATAGTGAAGGGAGAGATA 1413  
QY 1081 ATTCAAGATGCAAGAAAGGCCAATTTGATCTTCTGTAAGCTTATTTCTCTGAT 1140  
DB 1414 ATTCAAGATGCAAGAAAGGCCAATTTGATCTTCTGTAAGCTTATTTCTCTGAT 1473  
QY 1141 AGATCTGCACTTTTAATAGTCAATCTCAAGCAAAAACATATACATGGAACGATGTAC 1200

DB 1474 AGATCTGCACTTTTAATAGTCTCAAGCAAAAACATATACATGGAACGATGTAC 1533  
QY 1201 TCAGAGAAATGCTTCAAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTATC 1260  
DB 1534 TCAGAGAAATGCTTCAAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTATC 1593  
QY 1261 TCAGCCAGAGAACAAATGCCCAACTTTTAACTTTTAACTTTTAAAGAAAGATCCAGTAT 1320  
DB 1594 TCAGCCAGAGAACAAATGCCCAACTTTTAACTTTTAAAGAAAGATCCAGTAT 1653  
QY 1321 TCTGATCTTTTGAAGAGCTGATTAACAATCAAGAGCTCCCAATCATCTTTTGCAGGA 1380  
DB 1654 TCTGATCTTTTGAAGAGCTGATTAACAATCAAGAGCTCCCAATCATCTTTTGCAGGA 1713  
QY 1381 AAAAATCTTTTCCATTTTGAGACCCGACACTGAGCTGAACCCGTAACAAGTGTAT 1440  
DB 1714 AAAAATCTTTTCCATTTTGAGACCCGACACTGAGCTGAACCCGTAACAAGTGTAT 1773  
QY 1441 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTAGAAATAGACATCAAGTATC 1500  
DB 1774 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTAGAAATAGACATCAAGTATC 1833  
QY 1501 AACCGGACCTTCAGAGGCCATTCAGATTTGCAAGAGACACATCAAAAATGCTGAGT 1560  
DB 1834 AACCGGACCTTCAGAGGCCATTCAGATTTGCAAGAGACACATCAAAAATGCTGAGT 1893  
QY 1561 GATTAACAAAGTCAAAAAGAACTCTGATGCTTGTATTAATGCAATTTGTAAACAGCAA 1620  
DB 1894 GATTAACAAAGTCAAAAAGAACTCTGATGCTTGTATTAATGCAATTTGTAAACAGCAA 1953  
QY 1621 AATAACATGAATAATATATGATGCACTTCAAGATTAACCTGAGATTAATCAACAAGAAAT 1680  
DB 1954 AATAACATGAATAATATATGATGCACTTCAAGATTAACCTGAGATTAATCAACAAGAAAT 2013  
QY 1681 GTTTTGGCTCAGATCTCTTTTGTGAACAGAGCAAGACTAGGGGTATGAGCCACATAG 1740  
DB 2014 GTTTTGGCTCAGATCTCTTTTGTGAACAGAGCAAGACTAGGGGTATGAGCCACATAG 2073  
QY 1741 GGTATACGAATTCGTACATTTAAGAAAGCTTCAATCTCGTGTGTGCTCAACAGTTTAAA 1800  
DB 2074 GGTATACGAATTCGTACATTTAAGAAAGCTTCAATCTCGTGTGTGCTCAACAGTTTAAA 2133  
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCACTTCTGATTCAGAGAGTGTGT 1860  
DB 2134 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCACTTCTGATTCAGAGAGTGTGT 2193  
QY 1861 GTGAGCTTGTAAAGAGATGATCATCTCAGAAATATGTGAAAGAGTTCTTCAATATCT 1920  
DB 2194 GTGAGCTTGTAAAGAGATGATCATCTCAGAAATATGTGAAAGAGTTCTTCAATATCT 2253  
QY 1921 AGTATGGAATTAAGATCACTATTTATTTATTCAAATGCTGTATGAGTTTCCCTTGTCT 1980  
DB 2254 AGTATGGAATTAAGATCACTATTTATTTATTCAAATGCTGTATGAGTTTCCCTTGTCT 2313  
QY 1981 GATAGACACCCCTCACTACTGACCAACATCAGTAGTACAGCTTTGCAATTTTACAGAA 2040  
DB 2314 GATAGACACCCCTCACTACTGACCAACATCAGTAGTACAGCTTTGCAATTTTACAGAA 2373  
QY 2041 AAAATCTGGGAAATAATCAATATGCTTCCAGGTTTGTACAGCTTGTAAAGTCTAAATCT 2100  
DB 2374 AAAATCTGGGAAATAATCAATATGCTTCCAGGTTTGTACAGCTTGTAAAGTCTAAATCT 2433  
QY 2101 CCCAAAATCACTTTTATTAACAAGATATGCTAAATGCAATTTGATGAGAAATTCCTCGGT 2160  
DB 2434 CCCAAAATCACTTTTATTAACAAGATATGCTAAATGCAATTTGATGAGAAATTCCTCGGT 2493  
QY 2161 GCTGATTTTGAAGTTGTGTTTATGATGGGTAAATAACAACAACAAGAAATTCATT 2220  
DB 2494 GCTGATTTTGAAGTTGTGTTTATGATGGGTAAATAACAACAACAAGAAATTCATT 2553  
QY 2221 CAGGTGATTTGAAAAAGAGGGAAGTCTTACATTTTAAAAAGTGAAGTGAATATAGC 2280

Db 2554 CAGGTGATGTAAGAGCAGGAGAGCTTACCTTTAAAAAGTGAAGTAAATAGC 2613  
 QY 2281 TTGAAAAGAGAGATTAATAATGTATATGACCATGCTAATAGAGGTGATGTTT 2340  
 Db 2614 TTGAAAAGAGAGATTAATAATGTATATGACCATGCTAATAGAGGTGATGTTT 2673  
 QY 2341 GCATGGAATCCATAATTTTCAAGAGAGGAAAGAAACTAGAGTGTCTCTTTTCC 2400  
 Db 2674 GCATGGAATCCATAATTTTCAAGAGAGGAAAGAAACTAGAGTGTCTCTTTTCC 2733  
 QY 2401 ATATCATAGAGAGAAAGAACTGTAGTACTAGTTCCTAAGGCTTATCACTCTCT 2460  
 Db 2734 ATATCATAGAGAGAAAGAACTGTAGTACTAGTTCCTAAGGCTTATCACTCTCT 2793  
 QY 2461 TCTGTGATTTCAAAATTAACCAAGAGAGATAGCATCTTTCAACAGAAATGTCAT 2520  
 Db 2794 TCTGTGATTTCAAAATTAACCAAGAGAGATAGCATCTTTCAACAGAAATGTCAT 2853  
 QY 2521 AGTGTGCTTCTCCACACAGAGCAACCAATCTTAATCCCTATGTTACAAATGAAGA 2580  
 Db 2854 AGTGTGCTTCTCCACACAGAGCAACCAATCTTAATCCCTATGTTACAAATGAAGA 2913  
 QY 2581 CTGTGCTTCAACAATAAGCTTCTGAGAAAGACATCTCTTAATAGTCTAAAGATGT 2640  
 Db 2914 CTGTGCTTCAACAATAAGCTTCTGAGAAAGACATCTCTTAATAGTCTAAAGATGT 2973  
 QY 2641 CTTCCTAATTCAGACAACTTTTGAATCTGTTTGTGAAAATGTTGGTGGCTACA 2700  
 Db 2974 CTTCCTAATTCAGACAACTTTTGAATCTGTTTGTGAAAATGTTGGTGGCTACA 3033  
 QY 2701 CAGTTAAGTATGAGAGCTGTGTGAGTTCAGTTAATGATGGGCTCAGTTGTTGAC 2760  
 Db 3034 CAGTTAAGTATGAGAGCTGTGTGAGTTCAGTTAATGATGGGCTCAGTTGTTGAC 3093  
 QY 2761 GCAGAGGTCTTCTATCACTTATACCTCAACAAATGCTCAACAACTAGATGAGAA 2820  
 Db 3094 GCAGAGGTCTTCTATCACTTATACCTCAACAAATGCTCAACAACTAGATGAGAA 3153  
 QY 2821 AATGAAAATTAACAGATCAATCAACAAAGAAATTAAGTGTCTGTTCTCATCTTTG 2880  
 Db 3154 AATGAAAATTAACAGATCAATCAACAAAGAAATTAAGTGTCTGTTCTCATCTTTG 3213  
 QY 2881 ATGTTTCTAATCGACTCTCTAATTTTCAATGA 2913  
 Db 3214 ATGTTTCTAATCGACTCTCTAATTTTCAATGA 3246

# RESULT 2

US-08-252-995D-3  
 ; Sequence 3, Application US/08252995D  
 ; Patent No. 5650501  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dennie, James W  
 ; APPLICANT: Heffernan, Mike  
 ; APPLICANT: Fode, Carol  
 ; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BERESKIN & PARR  
 ; STREET: 40 King Street West  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5H 3Y2  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/252, 995D  
 ; FILING DATE: 02-JUN-1994  
 ; CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
 NAME: Kurdyak, Linda M  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-96  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3447 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Mus musculus  
 DEVELOPMENTAL STAGE: Lymphoid cDNA library  
 IMMEDIATE SOURCE:  
 LIBRARY: Murine Lymphoid  
 CLONE: MGA-resistant chop clones  
 FEATURE:  
 NAME/KEY: 5'UTR  
 LOCATION: 1..205  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 206..2980  
 FEATURE:  
 NAME/KEY: 3'UTR  
 LOCATION: 2981..3447  
 US-08-252-995D-3

Query Match 64.5%; Score 1879; DB 1; Length 3447;

Best Local Similarity 80.9%; Pred. No. 0; Mismatches 410; Indels 147; Gaps 9;

Matches 2362; Conservative 0;

QY 1 ATGCGACCTGATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTTGTAA 60  
 Db 206 ATGCGCGCTGATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTTGTAA 265  
 QY 61 GATCATTTTGTGTCTTCAAGAGTGAATTCATCTGATGTTGAAATGCAATC 120  
 Db 266 GATCATTTTGTGTCTTCAAGAGTGAATTCATCTGATGTTGAAATGCAATC 325  
 QY 121 AAAATGATGATTAAGAACCATGTAACAAGAGAGATGTAAGAGATCCAAATGAG 180  
 Db 326 AAAATGATGATTAAGAACCATGTAACAAGAGAGATGTAAGAGATCCAAATGAG 385  
 QY 181 GTGAATAATCATTTGCAATTTGAACATCTCTATCTTGAAGCTTATATTTTGA 240  
 Db 386 GTGAATAATCATTTGCAATTTGAACATCTCTATCTTGAAGCTTATATTTTGA 445  
 QY 241 GATAGCAATTAATGTATCTGTGATTAAGATGTGCAATTAAGAGAAATGAACAGAT 300  
 Db 446 GATAGCAATTAATGTATCTGTGATTAAGATGTGCAATTAAGAGAAATGAACAGAT 505  
 QY 301 CTAAAGAAATGAGTGAACCTTTCTCAAGAAATGAAGCTTCACATTCATGACCAATC 360  
 Db 505 CTAAAGAAATGAGTGAACCTTTCTCAAGAAATGAAGCTTCACATTCATGACCAATC 565  
 QY 361 ATCAACAGAGATGTGATCTTCAATCTCAATGATTAATCAACAGAGATCTCAACATTTCT 420  
 Db 566 ATCAACAGAGATGTGATCTTCAATCTCAATGATTAATCAACAGAGATCTCAACATTTCT 625  
 QY 421 AACCTCTAATGATCTGTAATTAAGACATCAAGATTTGAGTGTGAGTCAACTCA 480  
 Db 626 AACCTCTAATGATCTGTAATTAAGACATCAAGATTTGAGTGTGAGTCAACTCA 685  
 QY 481 CTGAAAATGCAATGAAAAGACATTAATTAATGAGAACTTCACTCAATTTTCA 540  
 Db 686 CTGAAAATGCAATGAAAAGACATTAATTAATGAGAACTTCACTCAATTTTCA 745  
 QY 541 GAAATGCACTGGAAGTCAATGAGCTTGAATCTGATGTTGATCTGAGGCTGATG 600

Db 746 GAAATTCGAACTGGAAGTGCACATGGAATTGATATTGGCTATTGGGCTGTAG 805  
Qy 601 TTTTATACATTACTTATGCGGAGACCACTCTTGACACTGACACAGTCAAGAACACTTA 660  
Db 806 TCTTATACGTTACTTATGGAAGACCACTTTTGACACTGACACAGTCAAGAACACTTA 865  
Qy 661 AATTAAGATGATGCGATTTATGAAATGCCATCTTTTGGTCATATAGAGGCAAGAG 720  
Db 866 AACAAAGATGCTGCGACATTTATGAAATGCCATCTTTTGGTCAGAGAGGCCAGAG 925  
Qy 721 GTTATTCACAGTACTTCTGTAGAAATTCAGACAGATCGTTTAAGTCTGTCTTCAATTTG 780  
Db 926 CTATTCACAGATTACTTGTAGAAACCTGCGAGATCGTTAAGTCTGTCTTGTGTTG 985  
Qy 781 GACCAATCTTTTATGTCGGAAATTTCTTCACAAAAAGTAAAGATTAGAACTGTGGA 840  
Db 986 GACCAATCTTTTATGTCGGAAATTTCTTCACAAAAAGTAAAGACTAGAGGACTAGAG 1045  
Qy 841 GACTCAATGATAGTGGGAGTGCACAAATTTCTACTGCAATTAACAGCTTCTTCACTACC 900  
Db 1046 GACTCAATGATAGTGGGAGTGCACAAATTTCTTCACAAAAAGTAAAGACTTCTGTATCC 1105  
Qy 901 AGTATAGTGTGATTTATTTGA CAAAAGAGA CTTTGTATGTCAGGCACTCCCAAT 960  
Db 1106 AGTTGAGTGGAGCTTACTTGAC --AGAAAGCTTTGTGTGTCAACCACTTCCAAAT 1162  
Qy 961 AAAATGACTGTATTTCCAAAGAAATAAAGTCACTGATTTTCTTCTTCAAGAGTGA 1020  
Db 1163 AAAATGACTGTATTTCCAAAGAAATAAATTAAGTCACTT --TTCTTCAAGAGTGA 1219  
Qy 1021 AACAGTTTATATCTAGTGGGGAAT-----CAAGAAACAGTAAATGTGGAAGGGA 1074  
Db 1220 AGTAAATTTTGTATCTCAATGGGGAATCCAGAAAGAAAGTAAATGTGGAAGGGA 1279  
Qy 1075 AAGATTAATCAAGATGCGAAGAAAGGCCACATTCGATCTTCGATCTTGTAGAGTTATCC 1134  
Db 1280 AAGATTAATCAAGATGCGAAGAAAGGCCAGTTCGATCTTCGATCTTGTAGAGTTATCC 1339  
Qy 1135 TCGATAGATCTGGGACTCTTAATAGTCAAGTCTCAAGCAAAAACATATCAATGAGACA 1194  
Db 1340 TCGATAGATCTGGGACTCTTAATAGTCAAGTCTCAAGCAAAAACATATCAAGAGACT 1396  
Qy 1195 TGTCACTCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAAATGAGAG 1254  
Db 1397 TGTCACTCAGTGAATGCTTTCAGAGGCTTGAAGATC----- 1434  
Qy 1255 AGTACTCA CCAAGACAAAGTCCAAATTTTAACTTTTAAAGAAAAGACATCC 1314  
Db 1435 ----- 1434  
Qy 1315 AGTATGTTCTGATCTTTGAAAAGCCTGATTAACAATCAAGCACTGCAATCATCTTGT 1374  
Db 1435 -----ACTGATGAAATCAACAAGTTCACATCATCTGT 1471  
Qy 1375 CCAGGAAAAATCTCTTTTCATTTGCGACCCGACACTCAGACTGAAAACCGTCAACAG 1434  
Db 1472 CTAGGAAAAATCTCTTTTCATTTGCGACCCGACACTCAGATGGAATGTGAACGAG 1531  
Qy 1435 TGTGTTGGGAATCTGCAATTAATGTCTCATTTTAAAGAAAACATCAATATGACAGATC 1494  
Db 1532 TGTGTTGGGAATCTGCAATTAATGTCTCATTTTAAAGAAAACATCAATATGACAGATC 1591  
Qy 1495 AGCCCAAAACCGGAGCTTTCAGAGGCGCATCGAATTTGCGAAGAGACATCAAAAATGCG 1554  
Db 1552 AGCCCAAAACCGGAGCTTTCAGAGGCGCATCGAATTTGCGAAGAGCTTTCGAAACGCT 1648  
Qy 1555 TGAATGATACAAAAGTCAAAAAGACTGATGCTTCTGATATGCAATTCGTGAAA 1614  
Db 1649 TGAATGATACAAAAGTCAAAAAGACTGATGCTTCTGATATGCAATTCGTGAAA 1708  
Qy 1615 CAGCAAAATACATGAAATATATGATGCTGCACTTCAAGTAAACCTGAGATATCAACAA 1674  
Db 1709 CAGCTGAGTGCATGAAATATATGATGCACTTCAAGTAAACCTGAGATATCAAGTGCACAG 1768

Qy 1675 GAATGTTTTTGGCTCAGATCTCTTCTGGAACAGAGCAAGACTAGGGGTATGAGCCA 1734  
Db 1769 GA-----GCCGGGCTCATCTCTATTTGAAACAAAGCAAAATGAAATGAGAGTGT 1822  
Qy 1735 CCATGGGGTTATCAAAATGTGATTTAAGAGCAATTACATCTCCGTGGTGTCTCAAG 1794  
Db 1823 ACACTGGGTATCAAAACCTTAAAGATATTAAGATCTCTGATTTGTCAAG 1882  
Qy 1795 TTTAAACCAATCAGACAGAAAACAAAAGGCTGTGTGAGCATCTGATTTCAAGAG 1854  
Db 1883 TTTAAACCAATCAGACAGAAAACAAAAGGCTGTGTGAGCATCTGATTTCAAGAG 1942  
Qy 1855 GTGTGTGGAGCTTGTAAAGAGTATGATCTCAAGATATGTAAGAAAGTTCTGAG 1914  
Db 1943 GTGTGTGGAGCTTGTAAAGAGTATGATCTCAAGATATGTAAGAAAGTTCTGAG 2002  
Qy 1915 ATATCTAGTATGAAATTCATCTATTTTATTTATCCAAATGTGTGAGAGTTTCT 1974  
Db 2003 ATATCTAGTATGAAATTCATCTATTTTATTTATCCGAAGATGGAAGAGCTTCT 2062  
Qy 1975 CTGCTGATGACCACTTCACTTCTGACAAATCATGAGTACAGCTTGAATTTA 2034  
Db 2063 CTGCTGATGACCACTTCTGCTTCTGACAAATCATGAGTACAGCTTGAATTTA 2122  
Qy 2035 CCAAGAAAATACTGGGAAAATATCAATATGCTTCAGGTTGTACAGCTTGAATCT 2094  
Db 2123 CCAAGAAAATACTGGGAAAATATCAATATGCTTCAGGTTGTATCAAGTATGAAATCT 2182  
Qy 2095 AAATCTCCAAATATCAATTTTATCAAGATATGCTTAAATGCAATTTTATGAGAAATCT 2154  
Db 2183 AAATCTCCAAATATCAATTTTATCAAGATATGCTTAAATGCAATTTTATGAGAAATCT 2242  
Qy 2155 CTTGCTGATTTTGAAGTTTGTATATGATGAGGATTAATTAACAAAACAGAGT 2214  
Db 2243 CTTGCTGATTTTGAAGTTTGTATATGATGAGGATTAATTAACAAAACAGAGT 2302  
Qy 2215 TTTCAATGAGTATGAAAGACAGAGAGTCTTCACTTAAAGTGAAGTGAAGT 2274  
Db 2303 TTTCAATGAGTATGAAAGACAGAGAGTCTTCACTTAAAGTGAAGTGAAGT 2362  
Qy 2275 AATAGCTGAAAAGAGATTAATTAATATGACCAATGCTAATAGGGTCAATCT 2334  
Db 2363 AATAGCTGAAAAGAGATTAATTAATATGACCAATGCTAATAGGGTCAATCT 2422  
Qy 2335 TGTATGACCTGGAATCCATATTTTCAAGAGGAAAGAAATCAGAGAGTCTCCCTT 2394  
Db 2423 TGTATGACCTGGAATCCATATTTTCAAGAGGAAAGAAAGTCAAGAGGAGTCTCCCTT 2482  
Qy 2395 TTCCCAATATCATAGAGAAACCTGTGTAGTACTAGTTCACTTAAGGCTTATCACT 2454  
Db 2483 TTCCCAATATCATAGAGAAACCTGTGTATATCTAGTTCACTTAAGGCTTATCACT 2542  
Qy 2455 CCTCTTCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2514  
Db 2543 CCTCTCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2599  
Qy 2515 ATGATATGCTGCTTCTTCCAAACAGAGCACTAATCTTAAATCCCTATAGTTCAAT 2574  
Db 2600 ATGATATGCTGCTTCTTCCCAACAGAGCACTAATCTTAAATCCCTATAGTTCAAT 2659  
Qy 2575 GAAGACTTGTCTTCAACACTCAGCTTCTGGAACAGACATCTCTTAAATAGTCAAAA 2634  
Db 2660 GAAGACTTGTCTTCAACACTCAGCTTCTGGAACAGACATCTCTTAAATAGTCAAAA 2708  
Qy 2635 GATGTCTTCTTAAATCAGACCAATTTTGAATCTGTTTTTGAAGAAATGTGTGTTG 2694  
Db 2709 ----GTCTTCTTAAATCAGACCTTTTGAATCTGTTTTTGAAGAAATGTGTGTTG 2764  
Qy 2695 GCTACACAGTTAATCAGAGAGTGTGTGCTTCAATTAAGTGGTCCAGTTGTT 2754  
Db 2765 GCTACACAGTTAATCAGAGAGTGTGTGCTTCAATTAAGTGGTCCAGTTGTT 2824

QY 2755 GTGACGAGCAGAGTGTCTTCTATCAGTTATACCTCACCAGAAATGCTCAAACTAGTAT 2814  
|||  
DB 2825 GTCCAGGCGAGGATATCTTCATCAGTTACATCATCAGATGCTCAGACCACTAGTAT 2884  
|||  
QY 2815 GGAGAAATGAAATTTACCAAGCTACATCAACAGAAATTTACAGTGTCTGTCTTCATC 2874  
|||  
DB 2885 GGAGAAATGAAATTTACCTGATACATCAACAGAAATTTACAGTGTCTGTCTTCATC 2944  
|||  
QY 2875 CTTTGAATGTTCTTAATCGAGCTCCTATTTTTCATTGA 2913  
|||  
DB 2945 CTTTGAATGTTCTTAATCGAGCTCCTATTTTTCATTGA 2983  
|||

## RESULT 3

US-08-834-108-3  
Sequence 3, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
APPLICANT: Rode, Carol  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurydyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
DEVELOPMENTAL STAGE: Lymphoid cDNA library  
IMMEDIATE SOURCE:  
LIBRARY: Murine Lymphoid  
CLONE: WGA-resistant chop clones  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..205  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 206..2980  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 2981..3447  
US-08-834-108-3

Query Match 64.5%; Score 1879; DB 2; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

QY 1 ATGGCAGCTTGCATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTTGTAA 60  
|||  
DB 206 ATGGCGCGTGATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTTGTAA 265  
|||  
QY 61 GATCATTTTGTGTGTCTCAAGAGTGTCCATTCACATGTTTGAAGTTGCAATC 120  
|||  
DB 266 GATCATTTTGTGTGTCTCAAGAGTGTCCATTCACATGTTTGAAGTTGCAATC 325  
|||  
QY 121 AAAATGATGATGAAGAACCGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
|||  
DB 326 AAAATGATGATGAAGAACCGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 385  
|||  
QY 181 GTGAAATATCATTTGCCAATTTGAAACATCTTCTATCTTGAAGCTTATTTTGA 240  
|||  
DB 386 GTGAAATATCATTTGCCAATTTGAAACATCTTCTATCTTGAAGCTTATTTTGA 445  
|||  
QY 241 GATAGCAATTAATGTATCTGTGTATTAAGAAATGTCCATTAATGAGAAATGA 300  
|||  
DB 446 GATAGCAATTAATGTATCTGTGTATTAAGAAATGTCCATTAATGAGAAATGA 505  
|||  
QY 301 CTAAAGAAATGAGTGAACCTTCTCAAGAAATGAAGCTCGACATTCATGACCAATC 360  
|||  
DB 506 CTAAAGAAATGAGTGAACCTTCTCAAGAAATGAAGCTCGACATTCATGACCAATC 565  
|||  
QY 361 ATCAGAGGATGTGTATCTTCTCATGATCTCAACCGGAGCTCAACATTTCT 420  
|||  
DB 566 ATCAGAGGATGTGTATCTTCTCATGATCTCAACCGGAGCTCAACATTTCT 625  
|||  
QY 421 AACCTCTTATCTGATCTGTAATATGAACATCAAGATTTGTATTTGGCTGGCACTCA 480  
|||  
DB 626 AACCTCTTATCTGATCTGTAATATGAACATCAAGATTTGTATTTGGCTGGCACTCA 685  
|||  
QY 481 CTGAAATATGCAATGAAGAAAGCACTATATATGGAATCTCTCACTCAATTTTCA 540  
|||  
DB 686 TTGAATATGCAATGAAGAAAGCACTATATATGGAATCTCTCACTCAATTTTCA 745  
|||  
QY 541 GAAATGCGCATGGAAGTGCATGAGCTTGAATCTGATGTTTGGCTGGCTGTATG 600  
|||  
DB 746 GAAATGCGCATGGAAGTGCATGAGCTTGAATCTGATGTTTGGCTGGCTGTATG 805  
|||  
QY 601 TTTTATACATTAATTAATCGGAGAACCACTTGCACATGACAGTCAAGAACATTA 660  
|||  
DB 806 TTTTATACATTAATTAATCGGAGAACCACTTGCACATGACAGTCAAGAACATTA 865  
|||  
QY 661 AATAAGTATGATGGAATTAATGAATGCCATTTTGTCAATGAAGGCGAAGAC 720  
|||  
DB 866 AATAAGTATGATGGAATTAATGAATGCCATTTTGTCAATGAAGGCGAAGAC 925  
|||  
QY 721 CTTATTCACAGTTACTTGTGGAATTCAGAGATCGTTAAGTGTCTTCAGATATG 780  
|||  
DB 926 CTTATTCACAGTTACTTGTGGAATTCAGAGATCGTTAAGTGTCTTCAGATATG 985  
|||  
QY 781 GACCATCTTTTATGTCGGAATTTCTCAACAAAGTGAAGATTAGAACTGTGAA 840  
|||  
DB 986 GACCATCTTTTATGTCGGAATTTCTCAACAAAGTGAAGATTAGAACTGTGAA 1045  
|||  
QY 841 GACTCAATGATGATGAGGACATGACAAATTTCTATGCAATTAACGCTTTCAGATAC 900  
|||  
DB 1046 GACTCAATGATGATGAGGACATGACAAATTTCTATGCAATTAACGCTTTCAGATAC 1105  
|||  
QY 901 AGTATAGTGTAGTATTTTGAACAAAGAGCTTTGATGTCAGGACCTCCCAAT 960  
|||  
DB 1106 AGTATAGTGTAGTATTTTGAACAAAGAGCTTTGATGTCAGGACCTCCCAAT 1162  
|||  
QY 961 AAAATGATGATTTTCAAGAAATTAAGTTCAACTGATTTTCTTCAAGAGATGA 1020  
|||  
DB 1163 AAAATGATGATTTTCAAGAAATTAAGTTCAACTGATTTTCTTCAAGAGATGA 1219  
|||  
QY 1021 AACAGTTTATATCTAGTGGGGAAT-----CAAGAAACAGTAAATGTGAAGGGA 1074  
|||  
DB 1220 AACAGTTTATATCTAGTGGGGAAT-----CAAGAAACAGTAAATGTGAAGGGA 1279  
|||

```

QY 1075 AAGTAAATCAAGATGCAAGAAAGGCCAATTCTGATACCTTGTAGAGTTATTC 1134
    |||||
DB 1280 AAGAGATTAAGAGATGCAAGAAAGAGCGCATTTCTGATACCTGCGCAGACTATCC 1339
    |||||
QY 1135 TGTGATAGATCTGGGCACTTCTAATAGTCAAGTCTCAAGCAAAAACATATCAATGACGA 1194
    |||||
DB 1340 TGTGATAGAGCCAGCCCTCTAA--TCAGTCTCGAGCAAAAACATCTCAGAGAACGT 1396
    |||||
QY 1195 TGTCACTCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAAAATGAAAG 1254
    |||||
DB 1397 TGTCACTCAGTAAATGCTTTCAAAGCCGTAGAAAGTC----- 1434
    |||||
QY 1255 AGGTACTCACCACAGACAACTAATGCCAATTTTAACTTCTTTAAAGAAAAGACATCC 1314
    |||||
DB 1435 ----- 1434
    |||||
QY 1315 AGTATGTTCTGATCTTTTAAAGACCTGATTAACATCAAGCACTCTCAATCATCTTGT 1374
    |||||
DB 1435 -----ACTGATGAAAATCAACAGATTCATCATCTATGT 1471
    |||||
QY 1375 CCAGGAAAACTCTTTTCCATTTGCAAGCCGACACTCAGACTGAAAACCGTACAACAG 1434
    |||||
DB 1472 CTAGGAAAACTCTTTTCAATTTGCAAGCAGACCTCAGATGGAATGATACAGCAG 1531
    |||||
QY 1435 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGAAAACTAAGTAATGACAGCATC 1494
    |||||
DB 1532 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGGAATTAATGACACACCGTT 1591
    |||||
QY 1495 AGCCCAACCGGGAATCTGCAAGGCGCATCGAGATTTGCAAGAGACACATCAAAAATGCG 1554
    |||||
DB 1552 ACCCAACAGAGATTTCCAGAGCTATCCAGATTTGC--AGAGACGTTTACGAAACGCT 1648
    |||||
QY 1555 TGGATGATACAAAAGTCAAAAAGAACTGATGCTTCTGATATGCACTTCTGTAATA 1614
    |||||
DB 1649 TGGATGATACAGAGAGCCAGAGAAATGCTGATATCTTCTGCAATTTGATGCTGAAG 1708
    |||||
QY 1615 CAGCAAAATACATGAATATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1674
    |||||
DB 1709 CAGCTGAGTGCATGAATATATGATGCACTATAGCTGAGGCTGATCCACAG 1768
    |||||
QY 1675 GATATGTTTTTGGCTCAATCTCTTTTGAACAGAGAACATGAGGCTATGAGCCA 1734
    |||||
DB 1769 GA-----GCCGAGCTTACATCTCTATTTGAAACAAAGAAATGAGATGAGTGG 1822
    |||||
QY 1735 CCATGGGCTTATCAGATGCTATTAAGAGCACTTACATCTCCGTTGGTGTCTCAGAG 1794
    |||||
DB 1823 AACTGGGCTTACAGAAACCTTAAAGATTAATCTCTCTGATGCTCAGACA 1882
    |||||
QY 1795 TTTAAACCAATCAGACAGAAAAACAAAAGGCTGTGTGAGCATTTGATTCAGAGAG 1854
    |||||
DB 1883 TTTAAAGCCAAATCAGACAGAAAAACAAAAGGCTGTGTGAGCATCTTGAATTCAGAGAG 1942
    |||||
QY 1855 GTGTGTGTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTGAAGAAGTCTTTCAG 1914
    |||||
DB 1943 GTGTGTGTGAGCTTGTGAAGAGTGTGTGAGAAATATGTGAAGAAGTCTTTCAG 2002
    |||||
QY 1915 ATATCTAGTGAAGAAATAGATCTATTTATTTATCCAAATGTGTGTGAGAGTTCCT 1974
    |||||
DB 2003 ATATGAGTGAAGAGATGATGATGATCTGTTTATTTACCGAAGAGTGAAGAGGCTTTCCT 2062
    |||||
QY 1975 CTGTCTGATAGACCACTCTCACTACTACTATTTATTTATCCAAATGTGTGTGAGAGTTC 2034
    |||||
DB 2063 CTGTCTGATAGACCTCTCTGCTACTGACACATCAGTAGAGTTCAGCTTTTGAATCTA 2122
    |||||
QY 2035 CCAGAAAAATCTGTGGGAAAAATATCAATATGCTTCCAGGTTGTACAGCTTGTGAAGATCT 2094
    |||||
DB 2123 CCAGAAAAATCTGTGGGAAAAATATCAATATGCTTCCAGGTTGTACAGCTTGTGAAGATCT 2182
    |||||
QY 2095 AAATCTCCCAAAATCACTATTTTACAAATATGCTTAATGCACTTTTGAAGAGATTC 2154
    |||||
DB 2183 AAAATCTCCCAAAATCACTATTTTACAAATATGCTTAATGCTTAATTTTGAAGAGATTC 2242
    |||||
QY 2155 CCTGTGTCTGATTTTGAAGGTTTGTGTTTATGATGGGGTAAATAACAAAAACAGAAAT 2214
    |||||

```

```

DB 2243 CCTGTGTCTGATTTGAAATTTGTTTATGATGAGCCAAATATCATTAACCTGAAAT 2302
    |||||
QY 2215 TTCAATCAGGATTTGAAAGACAGGAAAGCTTACCTTTTAAAGTGAAGT 2274
    |||||
DB 2303 TTAATTTACATATTTGAGAAAAACAGGATTTCTTTAATTTTAAATTAATTAATGAAGT 2362
    |||||
QY 2275 AATAGCTTGAAGAGAGATTAATAATGATATGACCAATGCTAATGAGGCTATCGTAT 2334
    |||||
DB 2363 ACCAGCTTGAAGAGAGATTAATAATGATATGACCAATGCTAATGAGGCTATCGTAT 2422
    |||||
QY 2335 TGTATGACCTGAGATCCATTAATTTTCAAGAGAGAAAGAACTAAGAGTCTCCCTT 2394
    |||||
DB 2423 TGTGTGTCACTGGAATCTGTAATCTGAGAGAGAAAGAGAGAGGCTTCTCATTC 2482
    |||||
QY 2395 TTCCCAATATCATAGAGAAAGAAACCTGATAGTACTAGTTCCTTAAGGCTTATCACT 2454
    |||||
DB 2483 TTCCCTAATATCTGAGAGAAAGAAACCTGATATCTAGTTCCTTAAGGCTTATCACT 2542
    |||||
QY 2455 CCTCTTCTGTGATTAATAATTAACCAACGAGAGATGAGCATCTTCAACAGAAATGCT 2514
    |||||
DB 2543 CCTC---CTGTGAGCCCAAGCTGTGTAAAGGAGAGAGGCGTCAAGAGCAGACTGAGC 2599
    |||||
QY 2515 ATGCAATGCTGCTCTTCTTCCACACAGGACCAATCTTAATCCCTATGCTTGAAT 2574
    |||||
DB 2600 GTGAATAGTCCGCTTTCCTCCACAGAGTCCCAAGACTCAGCTTCACTGTGACAT 2659
    |||||
QY 2575 GAAGGACTGTGCTTACAACTACAGCTTCTGGAACAGACATCTCTTCAATAGTCAAAA 2634
    |||||
DB 2660 GAAAGACTTGGCCACAGACGACTGACAGAGAACAGGCTCTTTCAA----- 2708
    |||||
QY 2635 GATGTCTTCTTAATCAGACCAACTTTTGAATCTGTTTGTGAATAATGTGTGTTG 2694
    |||||
DB 2709 -----GTCCTCTAATCTGCAAGCTTTTGAATCTGTTTGTGAATAATGTGTGTTG 2764
    |||||
QY 2695 GCTACACATTAATGATGAGAGCTGTGTGCTTCACTTAAATGATGGGTCCAGTGTGTT 2754
    |||||
DB 2765 GCTACACAGCTTAATGAGGAGAGCTGTGTGCTTCACTTAAATGATGGGTCCAGTGTGTT 2824
    |||||
QY 2755 GTGAGGAGAGAGTGTCTTCAATGATTAATACCTACCAAAATGTCAAACTAGATAT 2814
    |||||
DB 2825 GTCCAGGAGAGATTTCTTCACTGATTAATCAATCAAGATGTCAAGCACTAGATAT 2884
    |||||
QY 2815 GGAGAAAAATGAATAATTCAGACCTACATCAACAGAAATTAAGATGTCTGTCTTCATC 2874
    |||||
DB 2885 GGAGAAAAATGAATAATTCAGATTAATCAACAGAAATTAAGATGTCTGTCTTCATC 2944
    |||||
QY 2875 CTTTGAATGTTTCTAATCCAGCTCTTAATTTTCAATGA 2913
    |||||
DB 2945 CTTTGAATGTTTCTAATCCAGCTCTTAATTTTCAATGA 2983
    |||||

```

```

RESULT 4
US-08-252-995D-5
/ Sequence 5, Application US/08252995D
/ Patent No. 5650501
/ GENERAL INFORMATION:
/ APPLICANT: Dennis, James W
/ APPLICANT: Heffernan, Mike
/ TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: BERESKIN & PARR
/ STREET: 40 King Street West
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5H 3Y2
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
DEVELOPMENTAL STAGE: Lymphoid cDNA library  
IMMEDIATE SOURCE:  
LIBRARY: Murine Lymphoid  
CLONE: WGA-resistant chop clones  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..205  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 206..1597  
US-08-252-995D-5

Query Match 33.3%; Score 970.4; DB 1; Length 1600;  
Best Local Similarity 85.0%; Pred. No. 76-263;  
Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

QY 1 ATGGGCACTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCGCTGGTAA 60  
DB ATGGCGCGGTGATCGGGGAGAGATCGAGATTTTAAAGTTGGAATCGCTGGTAA 265  
QY 61 GGATCATTTGCTGCTGTCTACAGAGCTGATTCATTCACATGTTTGAAGTGAATC 120  
DB GGATCATTTGCTGCTGTCTACAGAGCTGATTCATTCACATGTTTGAAGTGAATC 325  
QY 121 AAAATGATGATTAAGAAAGCCATGTAACAAGCAGAAATGTTACAGAGTCCAAAATGAG 180  
DB AAAATGATGATTAAGAAAGCCATGTAACAAGCAGAAATGTTACAGAGTCCAAAATGAG 385  
QY 181 GTGAAAATGATTCGCAATTTGAAACATCTTCTATCTGAGAGTTTAACTATTTGAA 240  
DB GTGAAAATGATTCGCAATTTGAAACATCTTCTATCTGAGAGTTTAACTATTTGAA 445  
QY 241 GATAGCAATTATGTATCTGTATGATTAAGAAATGTCATTAATGAGAAATGACAGTAT 300  
DB GATAGCAATTATGTATCTGTATGATTAAGAAATGTCATTAATGAGAAATGACAGTAT 505  
QY 301 CTAAAGATAGAGTGAACCCCTTTCAGAAATGAAAGCTGACATTCATGACCAAGATC 360  
DB CTAAAGATAGAGTGAACCCCTTTCAGAAATGAAAGCTGACATTCATGACCAAGATC 565  
QY 361 ATCAAGAGGATGTTGATCTTCTATCTGATGATTAAGAAATGTCATTAATGAGAAAT 420  
DB ATCAAGAGGATGTTGATCTTCTATCTGATGATTAAGAAATGTCATTAATGAGAAAT 625  
QY 421 AACCTCTACTGCTGCTATATGATTAAGAAATGTCATTAATGAGAAATGTCATTAAT 480  
DB AACCTCTACTGCTGCTATATGATTAAGAAATGTCATTAATGAGAAATGTCATTAAT 685  
QY 481 CTGAAAATGCTACATGAAGACCTATTAATGAGAAATGTCATTAATGAGAAATGTC 540  
DB CTGAAAATGCTACATGAAGACCTATTAATGAGAAATGTCATTAATGAGAAATGTC 745

QY 541 GAAATTCACATCGAAGTGCATGATGCTGATGATGCTGCTGCTGCTGATG 600  
DB 746 GAAATTCACATCGAAGTGCATGATGCTGATGATGCTGCTGCTGCTGATG 805  
QY 601 TTTTATACATTAATTAATGAGAAAGCCCTTTCAGATGACATGACATGAAGACATTA 660  
DB TTTTATACATTAATTAATGAGAAAGCCCTTTCAGATGACATGACATGAAGACATTA 865  
QY 661 AATTAAGTATGATGAGCAATTAATGAAATGCAATCTTTTGTGATGAGGCAAGGAC 720  
DB AATTAAGTATGATGAGCAATTAATGAAATGCAATCTTTTGTGATGAGGCAAGGAC 925  
QY 721 CTATTCACCAATTAATTAATGAGAAATGCAAGATGCTTAAATGCTGCTGATTAATG 780  
DB CTATTCACCAATTAATTAATGAGAAATGCAAGATGCTTAAATGCTGCTGATTAATG 985  
QY 781 GACCATCTTTATGCTGCGAAATTTCTTCAAAAAGTAAAGATTTAGAACTGAGAA 840  
DB GACCATCTTTATGCTGCGAAATTTCTTCAAAAAGTAAAGATTTAGAACTGAGAA 1045  
QY 841 GACTCAATGATAGTGGGATGCAATTTCTATGCAATTAACAGCTTCTCAGTACC 900  
DB GACTCAATGATAGTGGGATGCAATTTCTATGCAATTAACAGCTTCTCAGTACC 1105  
QY 901 AGTATAGTGTATGATTTATTTGACAAAAGACCTTTGATGCTGACGACATCCCAAT 960  
DB AGTATAGTGTATGATTTATTTGACAAAAGACCTTTGATGCTGACGACATCCCAAT 1162  
QY 961 AAAATGCTGATTTTCAAAAAGTAAAGTCACTGATTTTCTTCAAGAGATGGA 1020  
DB AAAATGCTGATTTTCAAAAAGTAAAGTCACTGATTTTCTTCAAGAGATGGA 1219  
QY 1021 AACAGTTTATATCTAGTGGGAAAT-----CAAGAAACAGTAAATGTAAGGAGGA 1074  
DB AACAGTTTATATCTAGTGGGAAAT-----CAAGAAACAGTAAATGTAAGGAGGA 1279  
QY 1075 AGATTAATCAAGATGCAAGAAAGGACCAATTCATGATCTTGTAGAGCTTATTC 1134  
DB AGATTAATCAAGATGCAAGAAAGGACCAATTCATGATCTTGTAGAGCTTATTC 1339  
QY 1135 TCTGATGATCTGGACCTTCTAATAGTCAAGTCAAGCAAAAACATTAATCAATGAA 1194  
DB TCTGATGATCTGGACCTTCTAATAGTCAAGTCAAGCAAAAACATTAATCAATGAA 1396  
QY 1195 TGTCACTGACGAAATGCTTCAAGTCCAAAGATGACAGAGAGTGAAGAG 1254  
DB TGTCACTGACGAAATGCTTCAAGTCCAAAGATGACAGAGAGTGAAGAG 1453  
QY 1255 AGTACTCAACCAAGACATGACATTTTAACTTCTTAAAGAAAGACATCC 1314  
DB AGTACTCAACCAAGACATGACATTTTAACTTCTTAAAGAAAGACATCC 1512  
QY 1315 AGTATCTGATCTTTTGAAGACCTGATTAATCAATCAAGCACTT 1360  
DB AGTATCTGATCTTTTGAAGACCTGATTAATCAATCAAGCACTT 1558

RESULT 5  
US-08-834-108-5  
Sequence 5, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Helfertman, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSER: BERSKIN & PAR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada

```

? ZIP: MSH 3Y2
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/834,108
? FILING DATE:
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Kurdydk, Linda M
? REGISTRATION NUMBER: 34,971
? REFERENCE/DOCKET NUMBER: 3153-210
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 364-7311
? TELEFAX: (416) 361-1398
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1600 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Mus musculus
? DEVELOPMENTAL STAGE: Lymphoid cDNA library
? IMMEDIATE SOURCE:
? LIBRARY: Murine Lymphoid
? CLONE: MGA-resistant chop clones
? FEATURE:
? NAME/KEY: 5'UTR
? LOCATION: 1..205
? NAME/KEY: CDS
? LOCATION: 206..1597
? US-08-834-108-5

Query Match      33.3%; Score 970.4; DB 2; Length 1600;
Best Local Similarity 85.0%; Pred. No. 7e-263;
Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

```

```

QY 481 CTGAAATGCGCAGTAAAGCACTATATGAGAACTCTTACTTACCTTACCA 540
DB 686 TTGAATATGCGCAGTAAAGCACTATATGAGAACTCTTACTTACCTTACCA 745
QY 541 GAAATGCGCAGTAAAGCACTATATGAGAACTCTTACTTACCTTACCA 600
DB 746 GAAATGCGCAGTAAAGCACTATATGAGAACTCTTACTTACCTTACCA 805
QY 601 TTTTATCACTTATATGAGAACTCTTACTTACCTTACCTTACCA 660
DB 806 TCTTATCACTTATATGAGAACTCTTACTTACCTTACCTTACCA 865
QY 661 AATTAAGTATGAGAACTCTTACTTACCTTACCTTACCTTACCA 720
DB 866 AACAAAGTATGAGAACTCTTACTTACCTTACCTTACCTTACCA 925
QY 721 CTATATCACTTATATGAGAACTCTTACTTACCTTACCTTACCA 780
DB 926 CTATATCACTTATATGAGAACTCTTACTTACCTTACCTTACCA 985
QY 781 GACCATCTTATATGAGAACTCTTACTTACCTTACCTTACCA 840
DB 986 GACCATCTTATATGAGAACTCTTACTTACCTTACCTTACCA 1045
QY 841 GACTCAATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 900
DB 1046 GACTCAATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1105
QY 901 AGTATATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 960
DB 1106 AGTATATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1162
QY 961 AAAATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1020
DB 1163 AAAATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1219
QY 1021 AACAGTATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1074
DB 1220 AACAGTATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1279
QY 1075 AGAGTATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1134
DB 1280 AGAGTATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1339
QY 1135 TGTGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1194
DB 1340 TGTGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1396
QY 1195 TGTGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1254
DB 1397 TGTGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1453
QY 1255 AGGATATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1314
DB 1454 AGGATATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1512
QY 1315 AGTATATGATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1360
DB 1513 AATGATATAGTGGGATGCGCAGTAAAGCACTTATGAGAACTCTTACCA 1558

RESULT 6
US-08-252-995D-1
; Sequence 1, Application US/08252995D
; Patent No. 5650501
; GENERAL INFORMATION:
; APPLICANT: Dennis, James W
; APPLICANT: Heffernan, Mike
; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

```

ADDRESSER: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurd/dyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
DEVELOPMENTAL STAGE: Lymphoid cDNA Library  
IMMEDIATE SOURCE:  
LIBRARY: Murine Lymphoid  
CLONE: WGA-Resistant Chop Clones  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 206..1453  
NAME/KEY: 5'UTR  
LOCATION: 1..205  
US-08-252-995D-1

Query Match 31.7%; Score 924.2; DB 1; Length 1453;  
Best Local Similarity 86.4%; Pred. No. 6,6e-250;  
Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;

QY 1 ATGGCGACCTGCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
DB 206 ATGGCGCGCTGCATCGGGGAGAGATCGAGACTTTAAAGTTGGAATCTGCTGTAA 265  
QY 61 GGATCATTTCTGTGTCTTACAGAGCTGAGTCCATTCACTGTGTTGGAAGTTGCAATC 120  
DB 266 GGATCATTTCTGTGTCTTACAGAGCTGAGTCCATTCACTGTGTTGGAAGTTGCAATC 325  
QY 121 AAAAATGATGATTAAGAACCCATGTAAGAGAGAAAGTTGTAAGAGAGTCCAAAATGAG 180  
DB 326 AAAAATGATGATTAAGAACCCATGTAAGAGAGAAAGTTGTAAGAGAGTCCAAAATGAG 385  
QY 181 GTGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
DB 386 GTGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 445  
QY 241 GATAGCAATTTATGTGTATCTGTATGATTAAGAAATGTCATTAAGAGAAATGAAAGATAT 300  
DB 446 GATAGCAATTTATGTGTATCTGTATGATTAAGAAATGTCATTAAGAGAAATGAAAGATAT 505  
QY 301 CTAAAGATAGAGTGAAGACCTCTCAAGAAATGAAGCTGAGACATTCATGACCAAGATC 360  
DB 506 CTGAAGAAACGAATGAAGACCTTTCTCAAGAAAGGAAAGCTAGGACATTCATGACCAAGAT 565  
QY 361 ATCAGAGGATGTTGATCTTCAATTCATGATGATTAATCAACCGGAGCTCAGACTTTCT 420

DB 566 ATCAGAGGATGTTGATCTTCAATTCATGATGATTAATCAACCGGAGCTCAGACTTTCT 625  
QY 421 AACCTCTCTGATCTGTAATTAAGAACATCAAGATTTGATTTGGGCTGGCACTCA 480  
DB 626 AACATCTTAATTAAGGAGATTAAGAACATTAATTAAGGATTTGATTTGGGCTGGCACTCA 685  
QY 481 CTGAAAATGCCAATGAAGAAAGCACTTAATTAATGATGAACTCTTAATTAATTAATTAATTA 540  
DB 686 TTGAATATCCCAATGAAGAAAGCACTTAATTAATGATGAACTCTTAATTAATTAATTAATTA 745  
QY 541 GAAATTCGACCTGAAAGTGCACATGAGCTTGAATCTGATGTTGTTGCTGCTGGCTGTATG 600  
DB 746 GAAATTCGACCTGAAAGTGCACATGAGCTTGAATCTGATGTTGTTGCTGCTGGCTGTATG 805  
QY 601 TTTTATACATTAATTAATGAGAGACCACTTCAACATGACATGACATGACATGACATTA 660  
DB 806 TCTTATACATTAATTAATGAGAGACCACTTCAACATGACATGACATGACATGACATTA 865  
QY 661 AATTAAGATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCCCAAGAC 720  
DB 866 AACAAAGTATGCTGCGCAATTAATGAAATGCAATCTTTTGTCAATGAGCCCAAGAC 925  
QY 721 CTATTCACCAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 780  
DB 926 CTATTCACCAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 985  
QY 781 GACCAATCTTTTATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 840  
DB 986 GACCAATCTTTTATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1045  
QY 841 GACTCAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 900  
DB 1046 GACTCAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1105  
QY 901 AGTAATGATGATTAATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 960  
DB 1106 AGTTGATGAGGAGCTTCAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1162  
QY 961 AAAATGATGATTAATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1020  
DB 1163 AAAATGATGATTAATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1219  
QY 1021 AACAGTTTATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1074  
DB 1220 AAGTAATTTTATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1279  
QY 1075 AAGTAATTTTATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1134  
DB 1280 AAGTAATTTTATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1339  
QY 1135 TCTGATGATTAATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1194  
DB 1340 TCTGATGATTAATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1396  
QY 1195 TCTGATGATTAATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1253  
DB 1397 TCTGATGATTAATTAATTAATGAGAGATTAATGAAATGCAATCTTTTGTCAATGAGCTTCAATG 1453

RESULT 7  
US-08-834-108-1  
; Sequence 1, Application US/08834108  
; Patent No. 5976893  
GENERAL INFORMATION:  
; APPLICANT: Dennis, James W  
; APPLICANT: Helfferman, Mike  
; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERSKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto

STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdzyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
DEVELOPMENTAL STAGE: Lymphoid cDNA Library  
LIBRARY: Murine Lymphoid  
CLONE: MGA-Resistant Chop Clones  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 206..1453  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..205  
US-08-834-108-1

Query Match 31.7%; Score 924.2; DB 2; Length 1453;  
Best Local Similarity 86.4%; Pred. No. 6.6e-250;  
Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;

QY 1 ATGGGCACTGTCATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 60  
DB 206 ATGGGGCGTCGATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 265  
QY 61 GGATCACTTGGCTGCTGCTACAGAGCTGATTCATTCACACTGGTTTGAAGTTGCATC 120  
DB 266 GGATCACTTGGCTGCTGCTACAGAGCTGATTCATTCACACTGGTTTGAAGTTGCATC 325  
QY 121 AAAATGATGATTAAGAAAGCCATGTACAAGAGAGATGTACAGAGAGTCCAAATAG 180  
DB 326 AAAATGATGATTAAGAAAGCCATGTACAAGAGAGTGTACAGAGAGTCCAAATAG 385  
QY 181 GTGAATAATCATTTGCCAATTTAAACATCTTCTATCTTGGAGCTTTATTAATACTATTTGAA 240  
DB 386 GTGAATAATCATTTGCCAATTTAAACATCTTCTATCTTGGAGCTTTATTAATACTATTTGAA 445  
QY 241 GATAGCAATTTATGTGATCTGATTAAGAAATGTCGCAATTAAGGAATGAACAGTAT 300  
DB 446 GATAGCAATTTATGTGATCTGATTAAGAAATGTCGCAATTAAGGAATGAACAGTAT 505  
QY 301 CTAAAGATAGAGTAAACCTTCTCAGAAATAGAGCTCGACATTCATGACAGATC 360  
DB 506 CTGAAGAACAGATAGAGCTTCTTCAGAAAGAGAGCACTTATGACAGAT 565  
QY 361 ATCAAGAGATTTGATTTTATTTCTATCTCAATGTAATCAACCGGAGACTCACTTCT 420  
DB 566 ATCAAGAGATTTGATTTTATCTTCTATCTCAATGTAATCAACCGGAGACTCACTTCT 625  
QY 421 AACCTCTACTGACCTGTAATTAAGAAATCAAGATTTGCTGATTTTGGGCTGCAACTCA 480

DB 626 AACATTTTACTTACGGGAAATGAAATTAATTTGCTGACTTTGGACTTACAGACGAC 685  
QY 481 CTGAATAATCCCATGAAAGACATTAATATATGTAATCTTAATCACTTACATCA 540  
DB 686 TTGAATATGCCCATGAAAGACATTAATCACTGAGGACCTTAATATATTTACCA 745  
QY 541 GAAATGGCACTGAAAGGCAATGAGCTTGAATCTGAATGTTTGGTCCCTGGGCTGTAG 600  
DB 746 GAAATGGCACTGAAAGGCAATGAGCTTGAATCTGAATGTTTGGTCCCTGGGCTGTAG 805  
QY 601 TTTTATACATTTACTTATCGGAGACACCTTCGACACTGACACAGTCAAGAACATTA 660  
DB 806 TCTTATAGCTTACTTATGGAAGACACCTTTTGACATGACAGTCAAGAACATTA 865  
QY 661 AATTAAGTATGTTGGCAGATTAATGAATGCCATTTTGTCAATAGAGCCCAAGAC 720  
DB 866 AACAAAGTATGCTGAGATTAATGAATGCCAGCTTTTGTCAAGAGGCCCAAGAC 925  
QY 721 CTTATTCACAGTTACTTGTGAAATCCAGAGATGTTAAGTCTGCTCAATTTG 780  
DB 926 CTTATTCACAGTTACTTGTGAAATCCAGAGATGTTAAGTCTGCTCAATTTG 985  
QY 781 GACATTCCTTTATGTCGGAATTTCTTCAACAAAGTAAAGATTAGAACTGTGAA 840  
DB 986 GACATTCCTTTATGTCGGAATTTCTTCAACAAAGTAAAGATTAGAACTGTGAA 1045  
QY 841 GACTCAATGATAGTGGGCAATGTCATTTCTACTGCAATTAACGCTTTTCCAGTACC 900  
DB 1046 GACTCAATGATAGTGGGCAATGTCATTTCTACTGCAATTAACGCTTTTCCAGTACC 1105  
QY 901 AGTATAGTGTATGTTATTTTGAACAAAGAAAGACTTTGATTTGTCAGACACTCCCAAT 960  
DB 1106 AGTATAGTGTATGTTATTTTGAACAAAGAAAGACTTTGATTTGTCAGACACTCCCAAT 1162  
QY 961 AAAATGACTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTCAGAGATGGA 1020  
DB 1163 AAAATGACTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTCAGAGATGGA 1219  
QY 1021 AACAGTTTATTAATCACTGAGGGAAT-----CAGAAACAGTAATAGTGAAGGGA 1074  
DB 1220 AATAATTTTGTACTCAATGAGGGAATTCAGAAACAGTAATAGTGAAGGGA 1279  
QY 1075 AGAGTAATTAAGATGCAAGAAAGGCCATTTCTGATCTTCTGAGAGCTTAATCC 1134  
DB 1280 AGAGTAATTAAGATGCAAGAAAGGCCATTTCTGATCTTCTGAGAGCTTAATCC 1339  
QY 1135 TCTGATGATCTGGCACTTCTAATAGTCAATGTCAGAACAAACATTAATGGAACGA 1194  
DB 1340 TCTGATGATGCAAGCCCTCTTA--TCAGTCTGAGCAAAACATTAATGGAACGT 1396  
QY 1195 TGTCACTCAGAGAAATGCTTCAATGTCAGTCCAAAGATCA 1233  
DB 1397 TGTCACTCAGAGAAATGCTTCAATGTCAGTCCAAAGATCA 1435

RESULT 8  
US-09-016-434-649  
Sequence 649, Application US/09016434  
Patent No. 650038  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304



Db 1080 ACATCATTCGACATGACTTTT 1103

## RESULT 10

US-09-272-796-8  
Sequence 8, Application US/09272796  
Patent No. 6207148  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Guebler, Karl G.  
APPLICANT: Lal, Preeti  
APPLICANT: Goli, Surya K.  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/272,796  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/878,989  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0321 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2770 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HUVEH0801  
CLONE: 39043  
US-09-272-796-8

Query Match 4.6%; Score 134.4; DB 3; Length 2770;  
Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

Db 173 AAAATGAGTGAATAATACATTCGCAATGGAACATCCCTTATCTTGAGGCTTATTAAT 232  
Db 483 ACAAGAATAGAGCTTACAGAAATCTTCAATCATAGCAATGATGAGTTTACACT 542  
Qy 233 ATTTGAAGATGCAATTATGTATCTGTATTAGAAATGTGCCATATAGAGAAATGA 292  
Db 543 ACTTGAAGAGAAAGAAACATTTACATTTCTTGGAATACAGAGTGAAGTCAATG- 601  
Qy 293 ACAGGTACTTAAGAAATAGAGGAACCTTCTGAGAAATGAAGAGCTGACACTTATGC 352  
Db 602 --GCTCATATTTTGAAGCAAGAAAGGTGTGACAGACCCAGAGTTGATCTACTCA 659

Qy 353 ACAGATCATCAGAGGAGTGTGATCTTCATCTCATAGGTATACACACGGGAACTCA 412  
Db 660 GGCAGATTGTGTGAGCTGAATAATACCTTCATGAACAGAAATCTTGACAGAGATCTCA 719  
Qy 413 CACTTCTAACTCTCTACAGCTGACCTGTAAATGAAACATCAAGATTGCTGATTTGGCTGG 472  
Db 720 AACTAGGGAACCTTTTATTAATGAACCATGGAATGAATGAGGAGCTTGCTGG 779  
Qy 473 CAATCAATGAAATGCCACATGAAAGCAATATACATATGGAAGTCTTAATACATA 532  
Db 780 CAGCCAGCTAGAACCTTGGAAACAGAAAGAGAGAGATATGTGTATCCCAATATATC 839  
Qy 533 TTTCACAGAAATTTGCCATCTGAGTCAATGAGTCAATGAGTCTTGAATGTGTGCTCTGG 592  
Db 840 TCTCTCCGAAAGTCTTCAACAAAGAGATAGGCTGTGATCAGACATTTGGGCTCTGG 899  
Qy 593 GCTGTATGTTTATATTAATTAATCTTATGCGAGACCACTTCTGACATCTGACAGCTCAAG 652  
Db 900 GCTGTATGTTTATATTAATTAATCTTATGCGAGAGGCCCCCATTTGAAACTACAAATCTCAAG 959  
Qy 653 ACATCTTAATAAAGTGTATTTGGCAGATTAATGAATGCCATCTTTTGTCAATAGAG 712  
Db 960 AACTTATAGTGTAGTAAGGAGAGCAAGTATACAAATGCCCTCATTTGCTGGCTCTG 1019  
Qy 713 CCAAGACCTTATTCACCAAGTATCTTGTAGAAATCCAGAGATCTTAAATCTCTCTT 772  
Db 1020 CCAAGCACTTATTAATGCTATGTTGTCAAAACCCAGAGATCTGCCAGTTTGATG 1079  
Qy 773 CAGTATGACCATCTTTTATGT 796  
Db 1080 ACATCATTCGACATGACTTTT 1103

## RESULT 11

US-09-136-282-3  
Sequence 3, Application US/09136282  
Patent No. 6063609  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, KAREN  
APPLICANT: JACKSON, JEFFREY  
APPLICANT: HANSBURY, MICHAEL  
APPLICANT: NERURKAR, SANDHYA  
APPLICANT: KOSHAH, AMY  
APPLICANT: BOUZYK, MARK  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,282  
FILING DATE: 20-AUG-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,112  
FILING DATE: 20-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2789 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-136-282-3

Query Match 4.6%; Score 134.4; DB 3; Length 2789;  
Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```
QY 173 AAAATGAGTGAATAATACATTCGCAATGAAACATCTCTATCTTGAGCTTTAACT 232
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATAGCATAGTACGATTACACT 559
QY 233 ATTTGAGATGACATTTATGTGTATCTGTATTAAGAAATGSCATTAATGAGAAATGA 292
DB 560 ACTTGAGAGACAAAGAAACATTTACATCTCTTGGAATACAGTACAGAGGTCAATG- 618
QY 293 ACAGGATCTAAAGAAATAGAGTAAACCTTCTCAGAAATGAAAGCTCGACCTCATGC 352
DB 619 --GCTCATATTTTGAAGAGAAAGAGTGTGACAGAGCCAGAAAGTTCATCTACCTCA 676
QY 353 ACCAGATCATCAGAGGATGTTGTATCTTCATCTCATGTGTATACATCAACCGGACCTCA 412
DB 677 GGCAAGATGTGTCTGAGTGAATACCTTCATGAAACAGAAATCTTGACAGAGATCTCA 736
QY 413 CACTTCTAACCTCTCTACGACTCGTAATATGAACATCAAGATTGCTGATTTGGGCTGG 472
DB 737 AACTAGGAACTTTTATTAATGAAGCAGATGAAAGTTGGGAGCTTGGCTGG 796
QY 473 CAATCACTGAATAATGCCACATGAAAGACATATACATTAATGTGAACCTCTAATCA 532
DB 797 CAGCCAGGCTAGAACCTTTGAAACACAGAGAGAAAGATGTGTGATCCCAAAATTATC 856
QY 533 TTTCACGAGAAATGCGCATCGAGAGTGCATGSCCTTGAAATGTAATGTTGGTCCCTGG 592
DB 857 TCTCTCTAAGTCTCCAAACAAAGAGATGCTGTGATAGCATTTGGGCTGG 916
QY 593 GCTGTATGTTTATACATTAATCTTATCGGAGACCACTTGCACACTGACACAGTCAAG 652
DB 917 GCTGTATGTTTATACATTAATCTTATCGGAGAGCCCACTTGAATCAAAATCTCAAG 976
QY 653 AACATTAATAAAGATGATGAGTGAATGAAATGCAATGCTTTTGTCAATAGAG 712
DB 977 AACCTATAGTGTCATAGAGGAAAGGATATACATGCGTCTCATGCTGGCTCTG 1036
QY 713 CCAAGACCTTATATCCAGCTTACTCTGTAGAAATCCAGACAGATCGTTTAAGTCTCTT 772
DB 1037 CCAAGACCTTAAATGCTAGTATGTTGTCCAAAAACCAAGAGATGCTCCAGTTTGATG 1096
QY 773 CAGTATGAGCACTCTTTTATGT 796
DB 1097 ACATCATTCGACATGACTTTT 1120
```

RESULT 12  
US-09-505-744-3  
Sequence 3, Application US/09505744  
Patent No. 6245544  
GENERAL INFORMATION:  
APPLICANT: Karen M. Anderson  
APPLICANT: Mark M. Bouzyk  
APPLICANT: Michael J. Hanebury  
APPLICANT: Jeffrey R. Jackson  
APPLICANT: Sandhya S. Nerurkar  
APPLICANT: Amy K. Roehak  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
FILE REFERENCE: GH-70231-D1  
CURRENT APPLICATION NUMBER: US/09/505, 744  
CURRENT FILING DATE: 2000-02-16

EARLIER APPLICATION NUMBER: 09/136,282  
EARLIER FILING DATE: 1998-08-20  
EARLIER APPLICATION NUMBER: 60/056,112  
EARLIER FILING DATE: 1997-08-20  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 2789  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
US-09-505-744-3

Query Match 4.6%; Score 134.4; DB 3; Length 2789;  
Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```
QY 173 AAAATGAGTGAATAATACATTCGCAATGAAACATCTCTATCTTGAGCTTTAACT 232
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATAGCATAGTACGATTACACT 559
QY 233 ATTTGAGATGACATTTATGTGTATCTGTATTAAGAAATGSCCATTAATGAGAAATGA 292
DB 560 ACTTGAGAGACAAAGAAACATTTACATCTCTTGGAATACAGTACAGAGGTCAATG- 618
QY 293 ACAGGATCTAAAGAAATAGAGTAAACCTTCTCAGAAATGAAAGCTCGACCTCATGC 352
DB 619 --GCTCATATTTTGAAGAGAAAGAGTGTGACAGAGCCAGAAAGTTCATCTACCTCA 676
QY 353 ACCAGATCATCAGAGGATGTTGTATCTTCATCTCATGTGTATACATCAACCGGACCTCA 412
DB 677 GGCAAGATGTGTCTGAGTGAATACCTTCATGAAACAGAAATCTTGACAGAGATCTCA 736
QY 413 CACTTCTAACCTCTCTACGACTCGTAATATGAACATCAAGATTGCTGATTTGGGCTGG 472
DB 737 AACTAGGAACTTTTATTAATGAAGCCATGAAATGTTGGGAGCTTGGCTGG 796
QY 473 CAATCACTGAATAATGCCACATGAAAGACATATACATTAATGTGAACCTCTAATCA 532
DB 797 CAGCCAGGCTAGAACCTTTGAAACACAGAGAGAAAGATGTGTGATCCCAAAATTATC 856
QY 533 TTTCACGAGAAATGCGCATCGAGAGTGCATGSCCTTGAAATGTAATGTTGGTCCCTGG 592
DB 857 TCTCTCTAAGTCTCCAAACAAAGAGATGCTGTGATAGCATTTGGGCTGG 916
QY 593 GCTGTATGTTTATACATTAATCTTATCGGAGACCACTTGCACACTGACACAGTCAAG 652
DB 917 GCTGTATGTTTATACATTAATCTTATCGGAGAGCCCACTTGAATCAAAATCTCAAG 976
QY 653 AACATTAATAAAGATGATGAGTGAATGAAATGCAATGCTTTTGTCAATAGAG 712
DB 977 AACCTATAGTGTCATAGAGGAAAGGATATACATGCGTCTCATGCTGGCTCTG 1036
QY 713 CCAAGACCTTATATCCAGCTTACTCTGTAGAAATCCAGACAGATCGTTTAAGTCTCTT 772
DB 1037 CCAAGACCTTAAATGCTAGTATGTTGTCCAAAAACCAAGAGATGCTCCAGTTTGATG 1096
QY 773 CAGTATGAGCACTCTTTTATGT 796
DB 1097 ACATCATTCGACATGACTTTT 1120
```

RESULT 13  
US-09-136-282-1  
Sequence 1, Application US/09136282  
Patent No. 6063609  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, KAREN  
APPLICANT: JACKSON, JEFFREY  
APPLICANT: HANEUBURY, MICHAEL  
APPLICANT: NERURKAR, SANDHYA  
APPLICANT: ROEHAK, AMY  
APPLICANT: BOUZYK, MARK  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestlia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,282  
FILING DATE: 20-AUG-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,112  
FILING DATE: 20-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestlia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2783 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-136-282-1

Query Match 4.5%; Score 132.4; DB 3; Length 2783;

Best Local Similarity 51.9%; Pred. No. 6,3e-27;  
Matches 324; Conservative 0; Mismatches 296; Indels 4; Gaps 1;

173 AAAATGAGTGAATAATGATGCAATGCAATGGAACATCTTCTATCTTGAGCTTTATTAAT 232  
500 ACAAGAATAAGAGCTTCAAGAAATCTTCATCATTAAGCATGATGACATTTTACACT 559  
233 ATTTGAGATGAGCAATTATGATGATCTGATTTAGAAATGTCCTCAATAGAGAAATGA 292  
560 ACTTCGAGGACAAAGAAACATTTACATCTCTTGGAATCTGAGTGAAGGCAATGCG 619  
293 ACAGGTATCTAAGATAGAGTGAACCTTCTCAGAAATGGAAGCTGCACATTCATGCG 352  
620 TCAATTTTGAAG---CAAGAAAGGTGTTGACAGAGCCAGAAAGTTGATCTACTCA 675  
353 ACCAGATCATCAGAGGATGTTGATCTTCATCTCATGATTAATCAACCGGAGCTCA 412  
676 GGCAGATGTTGTCGACTGAATAATCCTTCATGAACAAAGAAATCTTGCAAGAGATCTCA 735  
413 CACTTTCTAACCCTCTACTGCTGTAATTAATGAACATCAAGATTCGATTTTGAGCTGG 472  
736 AACTAGGGAACCTTTTATTAATTAAGACCATGGAATTAAGTTGGGAGCTTCGCTCG 795  
473 CAATCTCAACTGAATAATGCAATGAAAAGCATATATATGGAATCTCTAATCA 532  
796 CAGCAGGCTGAAACCTTGAACACAGAAAGAAAGATATGATGATACCCCAATATATC 855  
533 TTTCACCAAGAAATGCGCATGGAAGTGCATGCGCTTGAATCTGATGTTGGTCCCTGG 592  
856 TCTCTCCTGAGAGTCTCTCAACAAACAAAGACATGCTGATGATCAAGCAATTTGGCCCTGG 915  
593 GCTGATCTTTTATACATTAATCTTAATCGGGAACCAACCTTGACAGCTGACAGATCAAGA 652  
916 GCTGTATATGATATCAATATGTTACTAGAGAGGCCCCCAATTTGAATCAATATCTCAAG 975

653 ACACATTAATAATGATAGTATGAGCAGATTAATGAAATGCCATCTTTTGTCAATAGAG 712  
976 AACTTATAGTGATTAAGGAAAGCAAGATATACAAAGCCGCTCTCATCTGAGCTCGT 1035  
713 CCAAGACCTTATTCACAGTACTTGTGTAATCCAGAGATCGTTTAAGTCTGCT 772  
1036 CCAAGCACTTATATGCTAGTATGTTGTCAAAACCAAGAGATAGGCTTATGATG 1095  
773 CAGTATGAGACATCTTTATGT 796  
1096 AATCATTCAGACATGACTTTT 1119

RESULT 14  
US-09-505-744-1  
Sequence 1, Application US/09505744

Patent No. 6245544  
GENERAL INFORMATION:  
APPLICANT: Karen M. Anderson  
APPLICANT: Mark M. Bouzyk  
APPLICANT: Michael J. Hanbury  
APPLICANT: Jeffrey R. Jackson  
APPLICANT: Sandhya S. Neturkar  
APPLICANT: Amy K. Roshak  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
FILE REFERENCE: GH-70231-D1  
CURRENT APPLICATION NUMBER: US/09/505,744  
CURRENT FILING DATE: 2000-02-16  
EARLIER APPLICATION NUMBER: 09/136,282  
EARLIER FILING DATE: 1998-08-20  
EARLIER APPLICATION NUMBER: 60/056,112  
EARLIER FILING DATE: 1997-08-20  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2783  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (2720) (2721)  
US-09-505-744-1

Query Match 4.5%; Score 132.4; DB 3; Length 2783;

Best Local Similarity 51.9%; Pred. No. 6,3e-27;  
Matches 324; Conservative 0; Mismatches 296; Indels 4; Gaps 1;

173 AAAATGAGTGAATAATGATGCAATGCAATGGAACATCTTCTATCTTGAGCTTTATTAAT 232  
500 ACAAGAATAAGAGCTTCAAGAAATCTTCATCATTAAGCATGATGACATTTTACACT 559  
233 ATTTGAGATGAGCAATTATGATGATCTGATTTAGAAATGTCCTCAATAGAGAAATGA 292  
560 ACTTCGAGGACAAAGAAACATTTACATCTCTTGGAATCTGAGTGAAGGCAATGCG 619  
293 ACAGGTATCTAAGATAGAGTGAACCTTCTCAGAAATGGAAGCTGCACATTCATGCG 352  
620 TCAATTTTGAAG---CAAGAAAGGTGTTGACAGAGCCAGAAAGTTGATCTACTCA 675  
353 ACCAGATCATCAGAGGATGTTGATCTTCATCTCATGATTAATCAACCGGAGCTCA 412  
676 GGCAGATGTTGTCGACTGAATAATCCTTCATGAACAAAGAAATCTTGCAAGAGATCTCA 735  
413 CACTTTCTAACCCTCTACTGACTGTAATTAATGAACATCAAGATTCGATTTTGAGCTGG 472  
736 AACTAGGGAACCTTTTATTAATTAAGACCATGGAATTAAGTTGGGAGCTTCGCTCG 795  
473 CAATCTCAACTGAATAATGCAATGAAAAGCATATATATGGAATCTCTAATCA 532  
796 CAGCAGGCTGAAACCTTGAACACAGAAAGAAAGATATGATGATACCCCAATATATC 855  
533 TTTCACCAAGAAATGCGCATGGAAGTGCATGCGCTTGAATCTGATGTTGGTCCCTGG 592

Db	856	TCTGCTCGAAGTCTCTCAACAAACAAAGACANTGGCTGTGAATTCAGACATTTGGGCTGTG	915
Qy	593	GCTGTATGTTTATATACATTACTTATTCGGGAACCACTTGGACACTGACACATGACAA	652
Db	916	GCTGTGAATGTATATACATGTCTACTAGGAGAGCCCCCATTTGAACTCAATCTCAAG	975
Qy	653	ACACATTTAAATAAGTAGTATTGGCAGATTATGAAATGCCATCTTTTGTGCAATAGAG	712
Qy	713	CCAAAGACCTTATTACACAGTTACTTGTGTAAGATCCAGACAGATGTTAAAGTGTCTT	772
Db	1036	CCAAGCACTTATTGCTAGTATGTGTCCAAAACCCAGAGATAGGCGTATGTTGATG	1095
Qy	773	CAGATTGAGACCATCTTTATGT	796
Db	1096	ACATCATTCGACATGACTTTT	1119
RESULT 15			
US-09-016-434-1147			
Sequence 1147, Application US/09016434			
Parent No. 6500938			
GENERAL INFORMATION:			
APPLICANT: Janice Au-Young			
APPLICANT: Jeffrey J. Seilhamer			
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING			
NUMBER OF SEQUENCES: 1490			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.			
STREET: 3174 PORTER DRIVE			
CITY: PALO ALTO			
STATE: CALIFORNIA			
COUNTRY: USA			
ZIP: 94304			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/016,434			
FILING DATE: HEREWITH			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER:			
FILING DATE:			
CLASSIFICATION:			
ATTORNEY/AGENT INFORMATION:			
NAME: Zeller, Karen J.			
REGISTRATION NUMBER: 37,071			
REFERENCE/DOCKET NUMBER: PA-0002 US			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (650) 855-0555			
TELEFAX: (650) 845-4166			
INFORMATION FOR SEQ ID NO: 1147:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 2169 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
IMMEDIATE SOURCE:			
LIBRARY: GENBANK			
CLONE: g1488262			
US-09-016-434-1147			

	Query Match	4.5%	Score 131.4;	DB 4;	Length 2169;
	Beech Local Similarity	50.1%;	Pred. No. 1.le-26;		
	Matches 355; Conservative	0;	Indels 3;	Gaps 1;	
Oy	50 TGCCTGGTAAGGATCATTTGCTGTGTCACAGCGTAGTCCATTGACACTGTTTTGG	109			

Db	119	IGTTGGGGAAGGGGGCTTTCGGCCGCTGTACGAGGCCACTGCACACAGAGACTGGCAGCG	178
Qy	110	AAGTTGCAATCAAAATGATAGATTAAGAAAGCCATGTACAAAGCAGAAATGATACAAAGAG	159
Db	179	CCTAAGCTGTCAAAATCATCTCCGACAGAGCCGGTGGCCCAAGCCGATCAGCCGGAAGA	238
Qy	170	TCGAAAATGAGGTGAATAATACATTTGCCAATTGAAAATCATCTCTTATCTTGAGCTTTATA	229
Db	239	TCCTTAATGAAATTAAGCTGCACCGAGACCTGCACACCCGCAATCTGTGCTTTTTCGC	298
Qy	230	ACTATTTTGAAGATGCAATTATGTGTATCTGTATTTAGAAAATGTCCATATATGAGAAA	289
Db	299	ACCACTTTGAGGACCGCTGACAACATCTCATATTTCTTGAGGCTGTGC---AGCCGAAAGT	355
Qy	290	TGAACAGGTATCTAAAGAAATAGAGTAAACCTTCTCAGAAAATGAAAGCTGCACCTTCA	349
Db	356	CCCTGGGCCCAATCTGGAAGGCCCGGACACACCTGTGGAGCCAGAAATGCGCTTACTACC	415
Qy	350	TGCACAGATCATCAACAGGAGATGTTGTATCTTTCATTCATGATATATAACAACCGGAGAC	409
Db	416	TGGCGCAGATCTTCTTGCGCCTCAAGTACTTGCACAGCGCGGCACTTGCACCGGGACC	475
Qy	410	TCACACTTCTTAACTCTCTACCTGCACTTCGTAATATGAAACATCAAGATGTGCTGATTTGGGC	469
Db	476	TCAGTTGGGAAATTTTTCATCACTGAGAAATGAAATGAACTGAAGGTGGGGATTTTGGGC	535
Qy	470	TGGCAATCAACTGAAAATGCCACATGAAAAGCATTATCATTTATGTGGAACTCTTAAC	529
Db	536	TGGCAGCCCGGTGGAGCGCTCCGGAGCAGAGAGAAAGAACATCTGTGTGCACCCCAACT	595
Qy	530	ACAATTCAACGAAATTTGCCACTCGAATGTCACATGCGCTTGAATCTGATGTTGGTCCC	589
Db	596	ATGTGTGCTTCCAGAAATGTGTGTGAACAAGGGCCAGCGCCCTTAAGGGATGTATGTGTAC	655
Qy	590	TGGGCTGATGTTTATATACATTACTTATCGGGAGACCAACCTTGCACATGACACAGTCA	649
Db	656	TGGGCTGTGTATGTACACGCTGCTCTGGGGAGCCCTTCTTGAACGGCTGACCTGA	715
Qy	650	AGAACACATTAATTAAGTATGATTTGGCAGATTATGAATGCCATCTTTTGTCAATAG	709
Db	716	AGGAGACGTACCGCTGCATCAAGCAGGTTCACTTACACGCTGCTGCACGCTCTCACTGC	775
Qy	710	AGGCCAAGACCTTATTCACACGTTACTTTCGTAAGAAATCCACACAGATCG	758
Db	776	CTGCCCGCAGTCTCTGGCCGCAATCTTCCGGGCTCAACCCGAGACCG	824

Search completed: September 24, 2004, 20:25:25  
Job time : 208 secs

**This Page Blank (usp10)**